Clinical and experimental studies on the role of APC in neoplasia

Cindy Anne Cooper

Doctor of Philosophy
The University of Edinburgh
2001

Declaration

The work presented in this thesis has been carried out by myself except where specifically stated within the text.

Cindy Cooper

2001

Acknowledgements

I would like to thank my PhD supervisors Professor Andrew Wyllie and Dr. Jill Bubb for their supervision throughout this project. My gratitude is also extended to Dr Alan Clarke for supervision during the transgenic research, and Owen Sansom for his continued advice on maintaining the transgenic colony. Finally, my thanks go to Robert Morris for advice on numerous technical aspects of this project and Dr Patricia Mote for proofreading the thesis at such short notice.

On a personal note, I would like to thank family and friends who have continuously supported and encouraged me, particularly during times of despair. Special thanks go to Chris who has waited patiently for me to complete this work.

Abstract

The adenomatous polyposis coli gene (*APC*) is located at human chromosome 5q21. Germline mutations in the *APC* gene characterise the Mendelian dominant inherited disorder, familial adenomatous polyposis coli (FAP). FAP patients develop numerous adenomas within the large intestine, some of which ultimately progress to carcinoma. A more general role for *APC* mutations in neoplasia is suggested by the fact that FAP patients are at risk of developing tumours of the brain, thyroid, bone and focal proliferative lesions ("desmoid tumours") of the connective tissue. Furthermore, mutations in the *APC* gene and / or loss of heterozygosity (LOH) at 5q21 have been reported in a range of sporadic tumours including tumours of the lung.

This thesis firstly investigated the occurrence of LOH at chromosome 5q21 in subtypes of human non-small cell lung cancer (NSCLC). Allele loss was compared in adenocarcinoma (ADC) and squamous cell carcinoma (SCC), the most common forms of NSCLC. Further sub-classification of the ADC subgroup was carried out to compare ADC arising in the bronchus versus the parenchyma, as determined by histology. Four genetic polymorphisms in the *APC* and *MCC* (mutated in colorectal cancer) genes at chromosome 5q21 were analysed for LOH in 92 consecutive primary SCC and ADC of the lung. Loss of heterozygosity was identified in at least two polymorphic loci in 40% (27/68) of informative cases. There was no significant difference in the frequency of LOH between SCC and ADC cases (Chi-squared test, df=1, p=0.446) or within the ADC subgroups; bronchial and parenchymal origin (Fisher's exact test; two-tailed test p=0.237). Results revealed that the frequency of LOH within the SCC group and the ADC remained similar at each tumour stage, Fisher's exact test (two-tailed test; p>0.2 in all cases). This result suggests that loss of 5q21 does not promote metastatic spread in these histological groups. Following the subdivision of ADC into site of origin, results revealed that LOH was not associated with increasing tumour stage in the parenchymal or bronchial subgroups, as determined by Fisher's exact test (two-tailed test; p>0.2 in all cases), or that the frequency of LOH differed between subgroups at each tumour stage,

Fisher's exact test, two tailed test; $p>0.2$ in all cases.

Thirty tumours showing LOH at one or more polymorphic sites were examined for mutations in the mutation cluster region (MCR) of *APC* by single-strand conformational polymorphism (SSCP) analysis. Mutations were not detected in any of these cases. These results suggest that it is likely that a tumour suppressor gene on 5q other than *APC* is involved in the pathogenesis of lung cancer.

Secondly, this thesis examined ADC of parenchymal or bronchial origin by the independent criterion of *K-RAS* mutation, a known feature in pulmonary ADC. Sixty-five surgically resected primary pulmonary adenocarcinomas were screened for mutations at codon 12 of the *K-RAS* gene by a PCR based method. Mutations in position 1 or position 2 of codon 12 were detected in 16 tumours (25%). When analysed by site of origin, mutations were seen in 9/26 (35%) parenchymal and in 0/12 bronchial ADC. This difference is significantly different, as determined by Fisher's exact test (two-tailed test; $p=0.0355$). No association was noted between the presence of *K-RAS* mutation at tumour stage, indicating that *K-RAS* mutations are not associated with metastatic spread. *K-RAS* mutations were also detected in 5 out of 32 foci of alveolar atypical hyperplasia (AAH), a lesion considered on histological grounds to be the precursor to parenchymal ADC. DNA sequencing showed that the great majority of mutations in both ADC and AAH were G-T transversions. These findings provide support for the classification of pulmonary ADC into bronchial and parenchymal subtypes and also provide molecular evidence to support the importance of AAH in the development of parenchymal cancers.

Finally, this thesis described the generation and initial characterisation of a murine *APC* transgenic founder line designed as a model to investigate the effects of aberrant expression of *APC*. Several *Apc* mutant murine models of FAP already exist. These all carry a heterozygous mutation in the *Apc* gene. The Min mouse (multiple intestinal neoplasia) is an example and carries a mutation at codon 850. All current models however are of limited use as mice homozygous for *Apc* mutations die at approximately

6.6 days post coitum, limiting analysis to *Apc* heterozygotes. Homozygous loss of *Apc* therefore depends upon additional somatic events that are not under direct experimental control and this may be associated with additional, undisclosed genetic events. Here a transgenic approach was taken to generate animals where the expression of the *APC* transgene is conditionally inactivated using the Cre-loxP recombination system of the bacteriophage P1. *APC* cDNA was flanked by loxP sequences and the promoter sequence of the ubiquitously expressed murine phosphoglycerate kinase (Pgk) gene cloned upstream of the *APC* cDNA to drive the expression of the transgene. Pro-nuclear injection was used to deliver the *APC* transgene into oocytes of wildtype mice (F_1 (C57BL/6 x CBA)). A transgene positive founder line was established and transcription of the *APC* cDNA confirmed in a wide variety of tissues using reverse transcription polymerase chain reaction (PCR). To date, transgene positive mice on an *Apc* wildtype background show no gross phenotype. Embryonic fibroblasts were derived and Cre-recombinase delivered by infection with a replication deficient adenovirus. *APC* cDNA excision was confirmed by PCR. The founder line was crossed with the C57BL/6 *Apc* $^{+/Min}$ line. Transgene positive *Apc* $^{+/Min}$ mice were interbred and offspring screened to identify whether the *APC* transgene can rescue the *Apc* $^{Min/Min}$ lethal phenotype. To date embryonic rescue has not been identified.

Figures

Abbreviations

A	-	adenine
AAH	-	alveolar atypical hyperplasia
ADC	-	adenocarcinoma
APC	-	adenomatous polyposis coli
ARF	-	alternative reading frame
BAC	-	bronchiolalveolar carcinoma
BS	-	brain specific
BSA	-	bovine serum albumin
C	-	cytosine
CAD	-	central contex dependent activation domain
CAK	-	CDK-activating kinase
CBP	-	CREB binding protein
CDI	-	CDK inhibitor
CDK	-	cyclin dependent kinase
cDNA	-	complementary DNA
CGH	-	comparative genomic hybridisation
CIN	-	chromosome instability
CIS	-	carcinoma *in situ*
CNS	-	central nervous system
CRD	-	cytosine rich extracellular domain
CTD	-	carboxyl terminal domain
DAB	-	diaminobenzidine
dATP	-	deoxyadenosine triphosphate
dCTP	-	deoxycytidine triphosphate
DDW	-	double distilled water
DEP	-	Dsh/egl-10/pleckstin
DEPC	-	diethylpyrocarbonate

dGTP	-	deoxyguanosine triphosphate
DHR	-	Dishevelled homologous region
DIX	-	Dishevelled and Axin
DNA	-	deoxyribonucleic acid
DNTP	-	deoxynucleoside triphosphate
Dpc	-	days post coitum
dTTP	-	deoxytyrosine triphosphate
EC	-	extracellular
EF	-	embryonic fibroblast
EGF	-	epidermal growth factor
EGFR	-	epidermal growth factor receptor
EMSA	-	electrophoresis mobility shift assay
ERK	-	extracellular signal regulated kinase
ES	-	embryonic stem
Fabp1	-	fatty acid binding protein gene
FAP	-	familial adenomatous polyposis coli
FRP	-	frizzled related proteins
FSH	-	follicle stimulating hormone
G	-	guanine
GAP	-	GTPase-activating protein
GDP	-	guanosine diphosphate
GEF	-	guanosine nucleotide exchange factor
GRB2	-	growth factor receptor bound protein 2
GSK-3β	-	glycogen synthase kinase
GTP	-	guanosine triphosphate
hCG	-	human chorionic gonadotrophin
HGF	-	hepatocyte growth factor
HLH	-	helix loop helix
HMG	-	high mobility group

kb	-	kilobase pairs
kDa	-	KiloDalton
L	-	Litre
LB	-	Luria-Bertani
LBD	-	ligand binding domain
LCC	-	large cell carcinoma
LOH	-	loss of heterozygosity
LoxP	-	locus of X-over of P1
MAGUK	-	membrane associated guanylate kinase
MAP	-	mitogen activated protein
MAPs	-	microtubule associated proteins
MB	-	Myc homology box
MCC	-	mutated in colorectal cancer
MCR	-	mutation cluster region
MEK	-	MAP extracellular signal regulated kinase
MIN	-	multiple intestinal neoplasia
MMP	-	matrix metalloproteinase
Mom	-	modifier of Min
mRNA	-	messenger RNA
MT	-	microtubule
NDSPT	-	normal donkey serum/PBS/Tween 20
NLS	-	nuclear localisation sequence
NMDA	-	N-methyl-D-aspartate
NSCLC	-	Non-small cell lung cancer
NTD	-	N (amino)-terminal domain
PBS	-	phosphate buffered saline
PCR	-	polymerase chain reaction
PGK	-	phosphoglycerate kinase
PI3K	-	phosphatidylinositol 3-kinase

PIP	-	phosphatidylinositol (4,5) diphosphate
PIP3	-	phosphatidylinositol (3,4,5) triphosphate
PKB	-	protein kinase B
PSD	-	post synaptic densities
PY	-	phosphotyrosine
RFLP	-	restriction length polymorphism
RGS	-	regulator of G protein signalling
RNA	-	ribonucleic acid
RT	-	room temperature
RTK	-	receptor tyrosine kinase
SAMP	-	amino acid sequence: serine, alanine, methionine, and proline
SCC	-	squamous cell carcinoma
SCLC	-	small cell lung cancer
SHC	-	Src homology and collagen
SH2	-	SRC homology domain 2
SSCP	-	Single strand conformational polymorphism
T	-	thymine
TBE	-	tris-borate-EDTA buffer
TGF	-	transforming growth factor
UTR	-	untranslated region
VNTR	-	variable tandem repeat
WHO	-	World health organisation
X-gal	-	5-bromo-4chloro-3-indolyl-β-D-galactoside
Zip	-	leucine zipper

Chapter 1: Introduction to the biology of non-small cell lung cancer

1.1 Epidemiology of lung cancer

1.1.1 Incidence

Lung cancer is the most common diagnosed cancer worldwide. High incidences of the disease occur in developed countries, particularly in North America and Europe, and are less common in developing countries, including Africa and South America (Parkin *et al.*, 1992; 1994). From an international perspective, Scotland has a comparatively high incidence and lung cancer is the commonest cancer in men and the second most common cancer in women (Parkin *et al.*, 1992; 1995; The Scottish Office, 1997). Statistics have shown that over a 10 year period (1986-1995) the incidence in males has decreased by approximately 15% whilst in females it has dramatically increased by 27% (The Scottish Office, 1997). This contrasting trend between the sexes is probably a result of changing smoking habits, with women adopting the habit later than men. The incidence of lung cancer peaks between 60 and 70 years of age and five-year survival is 7%, a figure that has remained relatively constant over the last 3 years (The Scottish Office, 1997; Scottish Health Statistics, 1998).

1.1.2 Smoking and lung cancer

Cigarette smoking is by far the leading cause of lung cancer, accounting for approximately 90% of lung cancer cases in countries where cigarette smoking is common, a fact that is reflected by the occurrence of lung cancer closely following trends in smoking (Peto 1994). The risk of lung cancer has been shown to increase with the duration of smoking and the number of cigarettes smoked per day (Doll and Peto 1978). Tobacco smoke contains thousands of components; to date 55 carcinogens have been identified and evaluated by the International Agency for Research on Cancer (Hoffmann

and Hoffmann 1997). The mechanisms by which tobacco smoke causes carcinogenesis is not fully understood although major advances have been made in understanding carcinogen activation by cellular detoxification processes and how free radicals and oxidants, contained within the smoke, result in lung damage (reviewed in Hecht 1999).

The major chemical classes of carcinogens in tobacco smoke are the nicotine-derived N-nitrosamines and polycyclic aromatic hydrocarbons (Hoffmann *et al.*, 1983). Both require metabolic activation to exert their carcinogenic effects.

During metabolic activation, carcinogens are enzymatically transformed to a series of metabolites in an attempt by the organ to convert them to forms that are more readily excreted. The cytochrome P450 enzymes encoded by the *CYP* family of genes usually carry out initial steps to oxygenate the substrate (Guengerich 1997). Other oxidative enzymes may be also involved but less frequently. The oxygenated intermediates formed may undergo further transformations by glutathione S-transferases, uridine-5-diphosphate-glucuronosyl transferases and other enzymes that detoxify the intermediate compound (Armstrong 1997; Burchell and Coughtrie 1997). The metabolic intermediates formed by the cytochrome P450s can form adducts with bases of DNA as well as yielding oxygen radicals which are also DNA damaging, producing DNA strand breaks in addition to modifications such as 8-hydroxyguanine, thymine glycol and 5-hydroxy methyl uracil (Kasai and Nishimura 1986; Imlay and Linn 1988).

Carcinogen metabolites commonly form DNA adducts at guanine or adenine nucleotides (Hecht and Hoffmann 1988; Dipple 1995). The cellular mechanisms for DNA repair; direct, base excision and nucleotide excision repair, are important in determining whether DNA adducts persist (Pegg *et al.*, 1995; reviewed in Sancar 1996; Singer and Hang 1997). If DNA adducts are not repaired this can result in the miscoding of DNA during DNA replication. Analysis of mutations present in tumour suppressor genes and oncogenes in lung cancers typically exhibit a high frequency of mutations at guanine nucleotides (discussed further in section 1.3).

Cigarette smoke also contains many oxidants and free radicals that are capable of initiating or promoting oxidative damage without the need for metabolic activation. One

such carcinogen, which is present in high concentrations, is nitric oxide. This oxidant can deaminate 5-methylcytosine residues of DNA to produce thymidine residues, resulting in nucleotide transition, cytosine to thymidine (Wink *et al.*, 1991).

1.1.3 Other hazards

1.1.3.1 Occupational exposure

Among cancers that have been associated with occupational exposures, cancer of the lung is the most common (Doll and Peto 1981). Several metallic compounds are proven lung carcinogens (Friberg and Cederlof 1978; Sunderman 1979); they include arsenic (Lee and Fraumeni 1969; Lloyd 1971), chromium (Alderson *et al.*, 1981) and nickel (Kreyberg 1978). An increased risk of lung cancer has been identified in coke oven workers who are exposed to tar and soot, both of which contain benzo(a)pyrene (Lloyd 1971). Miners are also at increased risk of lung cancer, possibly due to exposure to radon gas that leaks from igneous rocks (Radford and Renard 1984; Archer 1988).

Asbestos is well established as an occupation carcinogen (Selikoff *et al.*, 1976; 1980). The risk of lung cancer has been shown to increase with increased exposure. It is unclear whether asbestos acts directly as a carcinogen or through indirect mechanisms such as causing chronic inflammation (Brown 1986).

1.1.3.2 Diet

The link between diet and lung cancer, as with the majority of other cancers, is still poorly understood. In the majority of cases fruit and vegetables consumption has been observed to be inversely associated with lung cancer (Fraser *et al.*, 1991; Knekt *et al.*, 1991; Candelora *et al.*, 1992). Particular nutrients, retinol and carotene, and antioxidants within these foods have shown protective associations (Dartigues *et al.*, 1990; Cooper 1999). Dietary fat is thought to facilitate tumour growth but to date evidence correlating fat consumption and lung cancer risk have been inconsistent (Alavanja *et al.*, 1996; Nyberg *et al.*, 1998 a;b).

1.1.4 Genetic predisposition

Environmental agents and cigarette smoking have been shown to cause lung cancer in only a minority of people exposed. Epidemiology studies have demonstrated a family history of lung cancer predicts increased risk. These studies support a genetic basis for lung cancer susceptibility. Possible mechanisms that result in lung cancer susceptibility include polymorphisms in genes that result in differential rates of activation and elimination of respiratory carcinogens and polymorphisms within oncogenes and tumour suppressor genes.

1.1.4.1 Polymorphisms in genes involved in xenobiotic metabolism

Polymorphisms in genes encoding cytochrome P450 and glutathione s-transferase enzymes have been investigated for possible genetic susceptibility to lung cancer.

The cytochrome P450 enzymes, termed phase I enzymes, transform selected xenobiotics to intermediates that are detoxified by phase II enzymes which include the glutathione S-transferases. Many compounds in cigarette smoke e.g. polycyclic aromatic hydrocarbons and nitrosamines, are metabolised by the cytochrome P450 enzymes resulting in the production of intermediates that have carcinogenic effects (see section 1.1.2). Two cytochrome P450 enzymes, which are encoded by the genes termed *CYP1A1* and *CYP2D6,* have been investigated with respect to lung cancer risk as polymorphic variants have been shown to increase metabolic activation of carcinogens.

Evidence that polymorphisms within *CYP1A1* are associated with increased risk of lung cancer has varied across populations. A particular polymorphism has been associated with increased risk in the Japanese population but not Brazilian or American populations (Nakachi *et al.*, 1993; Shields *et al.*, 1993; Sugimura *et al.*, 1994). This difference between populations suggests that other genetic variations may be present which effect lung cancer susceptibility.

Genotypic variants of *CYP2D6* have been extensively investigated. This enzyme determines the phenotype for the metabolism of the anti-hypertensive drug debrisoquine. Individuals that are fast metabolisers of the drug were shown to have a 10 fold risk of

developing lung cancer (Gonzalez *et al.*, 1988; Gough *et al.*, 1990) conversely a case control study failed to support these findings (Shaw *et al.*, 1995). Analysis of different genotypes have also generated inconsistent results (Bouchardy *et al.*, 1996; Legrand *et al.*, 1996; Christensen *et al.*, 1997; London *et al.*, 1997).

The family of genes that encode glutathione S transferases (*GSTs*), phase II detoxifying enzymes, have also been extensively analysed. There are at least four genetically distinct classes of the glutathione S-transferases; *GSTM1, GSTM3, GSTP1* and *GSTT1*.

The presence of a polymorphic variant that encodes a deficient phenotype of *GSTM1* (present in 40% to 50% of the Caucasian population) reduces the efficiency of detoxification of carcinogens. Several groups have studied the distribution of the *GSTM1* null genotype to determine whether there is an association with susceptibility to smoking-related lung cancer but with contradictory results (McWilliams *et al.*, 1995; Nyberg *et al.*, 1998c). Polymorphisms in other family members of the *GST* genes, *GSTM3*, *GSTP1* and *GSTT1* have also been reported. Saarikoski and co-workers (1998) studied the distribution of different genotypes in lung cancer patients and population controls and found no statistically significant effect on lung cancer risk. However, when *GSTM3*, *GSTP3* and *GSTT1* genotypes were analysed in combinations that included *GSTM1* it was found that lack of *GSTMI* and *GSTT1* genes led to an increased susceptibility to squamous cell carcinoma. Increased susceptibility to smoking related lung cancer through the presence of null alleles of *GSTM1* and *GSTT1* has been corroborated by Kelsey and coworkers (1997).

Finally, N-acetyltransferases are involved in the activation and inactivation reactions of numerous xenobiotics. N-acetyltransferases are encoded by the *NAT1* and *NAT2* genes, of which there are several allelic variants that have different acetylation capacities. Reports have shown that *NAT1* genotypes have a significant association with smoking related lung cancer whilst no significant association has been found with *NAT2* (Bouchardy *et al.*, 1998; Abel-Rahman *et al.*, 1998).

1.1.4.2 Inherited polymorphisms in oncogenes and tumour suppressor genes

Polymorphic variants in well-characterised oncogenes and tumour suppressor genes have been implicated in lung cancer predisposition.

The p53 protein (see section 1.3.2.2) exhibits a polymorphism at amino acid 72, resulting in either a proline or an arginine residue at this position. A proline residue would form part of a proline rich domain, which is ablated in the p53 arginine form, these isoforms differ in their abilities to activate transcription and induce apoptosis (Thomas *et al.*, 1998). Some studies have shown increased risk of non-small cell lung cancer (NSCLC) in proline/proline homozygotes whilst other studies implicate arginine/arginine homozygotes. As with the analysis of this polymorphism in other cancers, no consensus has been reached regarding predisposition to NSCLC (Weston *et al.*, 1997; Tagawa *et al.*, 1998 Auer *et al.*, 1999).

Inheritance of lung cancer susceptibility due to inherited polymorphisms in oncogenes is exemplified by polymorphic variants in *H-RAS* and *L-MYC* genes. The presence of a rare *H-RAS* allele containing an insertion/deletion polymorphism has been found to confer a relative risk to NSCLC in Caucasian and African American populations (Heighway *et al.*, 1986; White *et al.*, 1990; Weston *et al.*, 1992; 1997). In the Japanese population a restriction fragment length polymorphism in the L-*Myc* gene (see section 1.3.1.2) has also been associated with NSCLC, this polymorphism has also been shown to be a marker for metastatic potential (Kawashima *et al.*, 1988; 1992; Weston 1997). Similar studies have been carried out on NSCLC samples from patients from Australia, Norway and North America (Fong *et al.*, 1995b) and no increased risk of lung cancer has been associated with this polymorphism in these populations.

1.2 Lung cancer classification

Bronchogenic carcinomas, malignant tumours of bronchial or bronchiolar epithelial origin, account for up to 95% of primary lung tumours. The remaining 5% are mostly bronchial carcinoids; these tumours may be malignant or benign and arise from mucous glands. Other tumours do occur, but are infrequent and are classified as miscellaneous tumours and include benign and malignant mesenchymal tumours, malignant lymphomas, lymphomatoid granulomatosis and hamartomas.

According to the classification published by the World Health Organisation (WHO) (1982) lung cancers are classified into four major histological types; squamous cell carcinoma, adenocarcinoma, small cell carcinoma and large cell carcinoma (see table 1). Other tumours are classified into groups known as rare tumour types (miscellaneous) and tumours consisting of more than one histological type.

For clinical, therapeutic and biological reasons, lung cancers are frequently classified into small cell (SCLC) and non-SCLC (NSCLC) groups. SCLC are rapidly growing, highly malignant tumours that arise within the central thorax. Usually the tumour has metastasised by the time of diagnosis therefore patients are often treated by chemotherapy. Five-year survival for patients with early stage tumours is approximately 20% whereas in cases where there is extensive disease survival is less than 1%. NSCLC account for 75-80% of lung cancers. In comparison to SCLC, NSCLC grow more slowly, metastasise locally and regionally before widespread dissemination. If diagnosed within the early stage, patients may undergo curative surgery. Patients with later stages of NSCLC are treated with chemotherapy although only 30-40% of tumours respond to this treatment. Survival is highly dependent on tumour staging, with a 5-year survival of 65% in early stages and 1-2% in the more advanced tumour stages. The currently used International Cancer Staging System (ISS) (Mountain, 1997) is shown in table 2.

This thesis addresses the genetics of NSCLC, which from this point will refer to the three most common histological types ADC, LCC and SCC, other tumour types will not be discussed further.

Table 1. Classification of bronchogenic primary lung tumours and incidences according to the WHO (Yesner *et al.*, 1982).

Squamous cell carcinoma	**(35-50%)**
Adenocarcinoma	**(15-35%)**
a) Bronchial derived acinar; papillary; solid carcinoma	
b) Bronchioloalveolar cell carcinoma	
Small cell carcinoma	**(20-25%)**
a) Oat cell (lymphocyte-like)	
b) Intermediate cell (polygonal)	
c) Combined (usually with squamous)	
Large cell carcinoma	**(10-15%)**
a) Undifferentiated; giant cell; clear cell	
Combined squamous cell carcinoma and adenocarcinoma	

Table 2: TNM lung cancer staging system (Mountain, 1997)

TUMOUR

TX: Demonstrable only by cytology of bronchopulmonary secretions
T1: Tumour <3cm without pleural or main stem bronchus involvement.
T2: Tumour >3cm or involvement of main stem bronchus 2cm from carnis, visceral, pleural or lobular atelectasis.
T3: Tumour with involvement of chest wall, diaphragm, mediastinal pleura, pericardium, main stem bronchus 2cm from carina, or entire lung atelectasis.
T4: Tumour with invasion of mediastinum, heart, great vessels, trachea, oesophagus, vertebral body, or carina or with a malignant pleural effusion.

NODE

N0: No demonstrable metastasis to regional lymph nodes
N1: Ipsilateral hilar or peribronchial nodal involvement
N2: Metastasis to contralateral mediastinal or hilar lymph nodes, ipsilateral or contralateral scalene, or supraclavicular lymph nodes.

Metastases

M0: No (known) distant metastasis
M1: Distant metastasis present

Using the above classification lung cancer stage is described as follows:

Stage grouping

Stage 1a	T1	N0	M0
Stage Ib	T2	N0	M0
Stage IIa	T1	N1	M0
Stage IIb	T2	N1	M0
	T3	N0	M0
Stage IIIa	T1-3	N2	M0
Stage IIIb	Any T	N3	M0
	T3	N1	M0
	T4	Any N	M0
Stage IV	Any T	Any N	M1

1.2.1 Histology of NSCLC

1.2.1.1 Squamous cell carcinoma

Squamous cell carcinomas (SCC) form bulky grey-white necrotic intra-luminal masses as well as invading peri-bronchial and parenchymal tissues and regional nodes. Histologically, like SCC arising elsewhere, they contain stratified scale-like cells, attached together by well-formed uniformly distributed desmosomes (also referred to as prickles or intercellular bridges). SCC is further characterised by keratinisation and both individual cell and epithelial pearl formation. SCC can range from well differentiated to poorly differentiated which is reflected by the extent of keratinisation and desmosome formation. In poorly differentiated SCC, the histology appears towards that of the undifferentiated large cell carcinoma (LCC) pattern. A histological example of a SCC is shown in figure 1a.

1.2.1.2 Adenocarcinoma

Adenocarcinoma (ADC) may arise from the major bronchi or occur as peripheral, subpleural tumours. They are firm, grey-white, and may have a mucoid cut surface. The characteristic morphological features of ADC include cuboidal or columnar cells with prominent microvilli, forming gland-like structures. Intracellular mucin granules are present and mucin may be secreted, papillary structures may also be present. As with SCC the degree of differentiation varies considerably, with areas of poorly differentiated ADC often indistinguishable from LCC. Adenocarcinomas can be sub-classified into four groups according to their histological characteristics; acinar, papillary, and solid with mucin production and bronchioloalveolar carcinomas (WHO 1982). The former 3 groups are often termed bronchial derived adenocarcinomas, which is misleading as the majority arise in the periphery of the lung. Histological variability is described by the 3 subgroups. The bronchioloalveolar carcinomas (BAC) are a distinctive subtype which are thought to arise from the terminal bronchioles and alveoli. These tumours are diffuse and tend to

follow pre-existing alveolar structures, termed lepidic growth. Histologically BAC consist of tall, columnar to cuboid epithelial cells that line up along the alveolar septa and project into the alveolar spaces in numerous branching papillary formations. Glandular structures are rare. BAC are further classified into 3 groups: mucinous, non-mucinous and sclerosing (Clayton, 1988). Sclerosing BAC is thought to represent an advanced stage of non-mucinous BAC (Barsky *et al.*, 1986).

Even though the WHO classification is widely used, it has been criticised as poorly reproducible due to cell heterogeneity. In 1987, Edwards proposed a new classification for ADC based on cellular morphology and site. ADC were categorised as parenchymal (arising in the periphery of the lung from functional tissue rather than the supporting or connective tissue) and bronchial (arising at the hilum). Parenchymal tumours contain cells characteristic of goblet cells, Clara cells, type II pneumocytes and ciliated bronchiolar lining cells and occasionally neuroendocrine cells whilst ADC arising in the bronchi contain bronchial epithelial cells or bronchi glands (Bolen and Thorning, 1982; Raino, 1983; Singh *et al.*, 1984). A histological example of an ADC arising in the parenchymal is shown in figure 1 (b).

1.2.1.3 Large cell carcinoma

Large cell carcinoma (LCC) frequently present as large bulky peripheral tumours. These carcinomas contain cells with no differentiation pattern, they have large and vesicular nuclei with prominent nucleoli and abundant cytoplasm, which can be either clear or granular (figure 1c). Some carcinomas contain large multinucleated cells and are classified as giant cell tumours (Nash and Stout, 1958), others with clear cytoplasm are termed clear cell carcinomas. It is thought that a number of LCC represent poorly differentiated SCC, ADC or SCLC as when cultured or xenografted most LCC developed differentiated features, especially those of adenocarcinoma. Some tumours remained undifferentiated and may represent a truly separate class of undifferentiated tumours (Gazdar and McDowell, 1988).

Figure 1: Three histological subtypes of non-small cell lung cancer

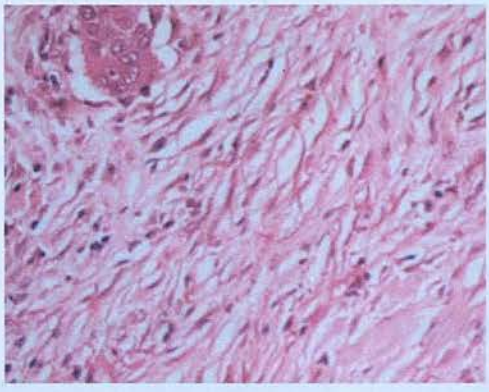
a) Haematoxylin and eosin (H+E) stained section of a squamous cell carcinoma. The characteristic feature of cellular stratification can be observed in this figure. Magnification x 400.

b) H+E stained section of a parenchymal adenocarcinoma. The arrow indicates the typical gland-forming characteristic of adenocarcinomas. Magnification x 400

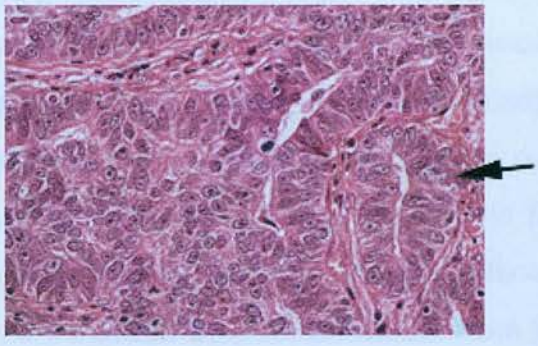
c) H+E stained section of a large cell carcinoma. Histological features include large vesicular nuclei with prominent nucleoli and abundant cytoplasm. Magnification x 400.

Figure 1 : Three histological subtypes of non-small cell lung cancer

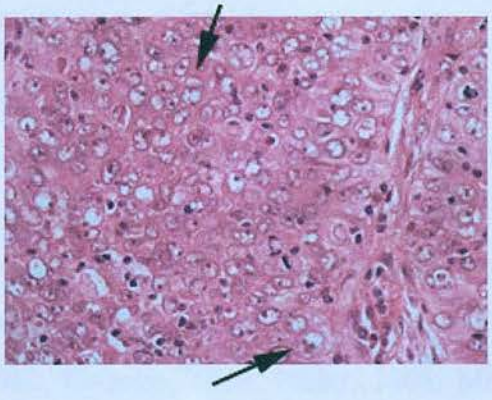
a)



b)



c)



1.2.2 Premalignant lesions of lung adenocarcinoma

Morphological analysis has identified several stages of premalignant lesions in the development of SCC. The spectrum of histological changes in preneoplastic bronchial epithelium extends from basal cell hyperplasia through squamous cell metaplasia to mild, moderate and marked dysplasia, carcinoma *in situ* and finally invasive cancer (Auerbach, 1957; 1961; Nasiell *et al.*, 1987). In contrast to SCC, the premalignant lesions for ADC have been less well characterised.

1.2.2.1 Pulmonary scars

Some investigators proposed pulmonary scars, which are frequently associated with ADC, as possible precursor lesions, however evidence is conflicting (Shimosato *et al.*, 1982; Madri and Carter, 1984; Kung *et al.*, 1985). Scars, identified by alveolar septate thickening and interstitial fibrosis, are often considered to be a result of healed tuberculous foci, healed or healing infarcts or due to pneumonia. The epithelium of small airways enclosed in the scar is thought to be particularly susceptible to malignant transformation (Madri and Carter, 1984). Shimosato and co-workers (1982) proposed that what had previously been regarded as a scar was in fact stroma induced by the growth of the ADC and suggested that ADC develop without any association with pre-existing scar tissue. Evidence to support this comes from the fact that in early-stage ADC, fibrotic foci are rarely detected.

1.2.2.2 Alveolar atypical hyperplasia

The best-known potentially premalignant lesions identified in the parenchyma of the lung are alveolar atypical hyperplastic lesions (AAH). These lesions have several synonyms; atypical epithelial proliferation, atypical alveolar cuboidal cell hyperplasia, bronchiolalveolar cell adenoma, alveolar adenomatous hyperplasia, atypical bronchiolalveolar cell hyperplasia (Raeburn and Spencer, 1953; Meyer and Liebow, 1965; Shimosato *et al.*, 1982; Kodama, 1986; Miller, 1988; Miller *et al.*, 1990; Weng *et al.*, 1990; Carey *et al.*, 1992). The term alveolar atypical hyperplasia or alveolar atypical

hyperplastic lesion (AAH) will be used from now on to describe these lesions.

Unlike areas of alveolar hyperplasia, which are identified as areas of proliferation of type 2 pneumocytes forming as a response to lung injury or part of the reparative process of the lung, AAH lesions are characterised by progressive cellular atypia of type II pneumocytes. Lesions are composed of a single row of atypical cuboidal cells with hyperchromatic nuclei and prominent nucleoli which line the alveolar walls in the absence of chronic inflammation (Nakanishi 1990; Carey *et al.*, 1992; Rao and Fraire 1995). The alveolar septa lined by the atypical cells are often thickened by fibrosis (Figure 2). AAH lesions have been classified into three grades; low grade, high grade and AAH-like carcinoma (Kitamura *et al.*, 1995;1996). Low-grade lesions are characterised by a single layer of round to cuboidal cells with low cellular density, small cell size, minimal variation in nuclear size and shape, and minimal thickening of the alveolar septa. In comparison, AAH lesions classified as high grade have increased cellular density, larger cell size, greater variation in cell size, shape, and mild fibrosis of the alveolar septa. The term AAH-like carcinoma is used to describe lesions that appear to represent a morphological transition between high-grade AAH and overt ADC. Morphological characteristics of this category consisted of lesions with high cellular density with stratification of epithelial cells, marked cytological atypia, and extensive thickening and fibrosis of the alveolar septa.

AAH lesions tend to be incidental findings following lung resections from patients with primary ADC. The reported frequency of AAH foci from patients resected for ADC has varied from 5.7-25% (Miller, 1990, Weng *et al.*, 1992; Carey *et al.*, 1992). Many lesions can be detected by gross examination of lung resections inflated by formalin or Bouin's fluid. Since AAH lesions are more difficult to identify grossly in routine specimens fixed by immersion, an underestimate of frequency is likely (Miller 1993), furthermore differential diagnosis between AAH and early invasive ADC with metastatic potential, such as BAC and papillary carcinomas with lepidic growth is extremely difficult.

The histological analysis of ADC and AAH lesions by Miller and co-workers (1988; 1990) led to the proposal that pulmonary ADC could follow a multistep progression from

precursor lesion to adenoma to carcinoma. Further to this, measurements of nuclear size as well as DNA content suggest a continuum from reactive hyperplastic type II cells to AAH and finally to ADC (Kodama *et al.*, 1986; Nakanishi 1990; Nakayama *et al.*, 1990; Kitamura *et al.*, 1995;1996; Mori *et al.*, 1996). To date there is no clear evidence that AAH lesions represent a neoplastic process rather than merely a reactive proliferation of non-neoplastic cells or that these lesions may be precursors of parenchymal ADC. **The experimental work described in this thesis further defines AAH lesions through analysis of the proto-oncogene *K-RAS*, a gene that is mutated in a high portion of ADC.**

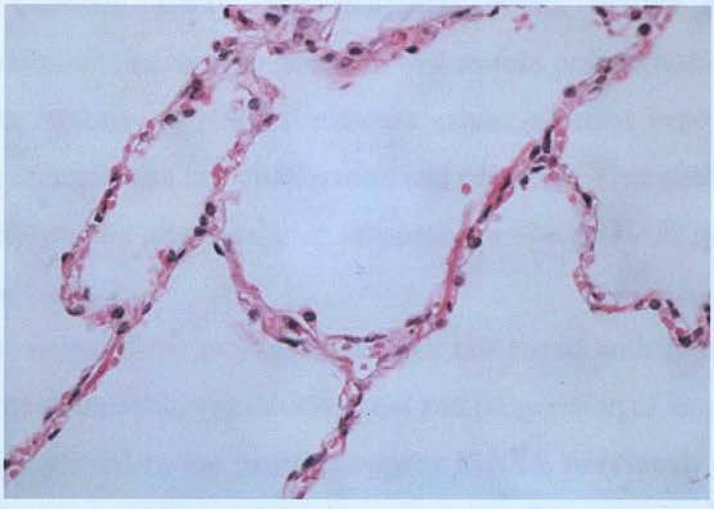
Figure 2: Histology of normal alveolar and an area of alveolar atypical hyperplasia

a) H+E stained section of alveolar in the normal lung. Magnification x 400

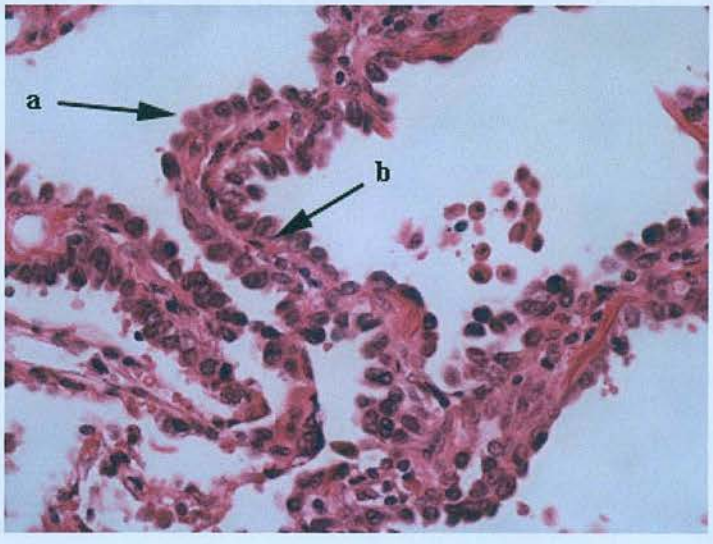
b) H+E stained section of a focus of alveolar atypical hyperplasia (AAH). A single row of hyperplastic type II pneumocytes (arrow a) line the thickened alveolar septa (arrow b). Magnification x 400

Figure 2: Histology of normal alveolar and an area of alveolar atypical hyperplasia

a)



b)



1.3 Molecular changes associated with NSCLC

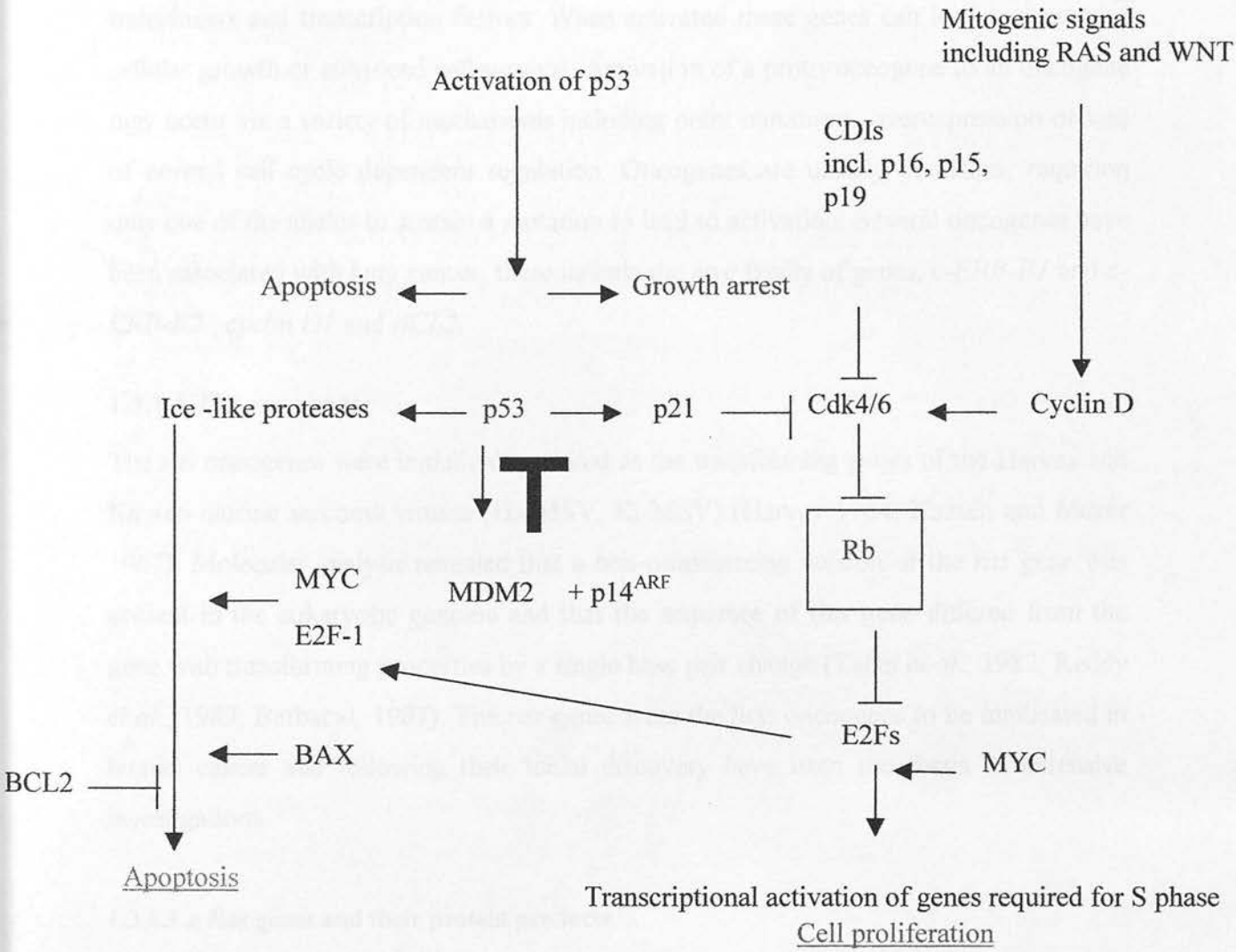
Lung carcinogenesis is a multistage process in which, by the time of clinical diagnosis, it is estimated that between 10 and 20 genetic alterations may have accumulated (Minna, 1993; Gazdar and Carbone, 1994). The molecular events include activation or overexpression of dominantly acting oncogenes as well as loss or inactivation of tumour suppressor genes. The majority of NSCLC tumours exhibit aberrant expression of key genes involved in the regulation of cell proliferation and cell death. One pathway that has been elucidated in controlling these cellular responses is the *p53-RB* pathway. This pathway is depicted in figure 3.

This section describes some of the molecular changes associated with primary NSCLC and relates these events to the multistep development and progression of lung tumours. **Particular emphasis is placed on the proto-oncogene *K-RAS*, previously shown to be mutated in NSCLC, as experimental work completed for this thesis addresses activation of this oncogene in subtypes of NSCLC and in alveolar atypical hyperplastic lesions. Furthermore, loss of heterozygosity at the chromosome region 5q21, the site of the tumour suppressor gene *APC*, is also of particular interest as I investigate a possible role of *APC* in NSCLC.**

Figure 3: The p53-RB pathway

The schematic representation in this figure outlines the regulation of cell proliferation, growth arrest and apoptosis through the interaction of a number of oncogenes and tumour suppressor genes. A further feature of this pathway is induction of *GADD45* by p53 to initiate DNA repair, this is not represented in this figure. The interaction of these proteins and subsequent consequences of aberrant expression is discussed within the text of chapter 2. The analysis of NSCLC has determined that aberrant expression of protein/or proteins within this pathway is a characteristic of these tumours.

Figure 3: The p53-Rb pathway



1.3.1 Proto-oncogenes and lung cancer

Proto-oncogenes are often growth factors, growth factor receptors, intracellular signal transducers and transcription factors. When activated these genes can lead to increased cellular growth or enhanced cell survival. Activation of a proto-oncogene to an oncogene may occur via a variety of mechanisms including point mutations, overexpression or loss of normal cell cycle dependent regulation. Oncogenes are usually dominant, requiring only one of the alleles to sustain a mutation to lead to activation. Several oncogenes have been associated with lung cancer, these include the *myc* family of genes, c-*ERB-B1* and *c-ERB-B2* , *cyclin D1* and *BCL2*.

1.3.1.1 The *ras* genes

The *ras* oncogenes were initially discovered as the transforming genes of the Harvey and Kirsten murine sarcoma viruses (Ha-MSV, Ki-MSV) (Harvey 1964; Kirsten and Mayer 1967). Molecular analysis revealed that a non-transforming version of the *ras* gene was present in the eukaryotic genome and that the sequence of this gene differed from the gene with transforming properties by a single base pair change (Tabin *et al.*, 1982; Reddy *et al.*, 1982; Barbacid, 1987). The *ras* genes were the first oncogenes to be implicated in human cancer and following their initial discovery have been the focus of extensive investigations

1.3.1.1.a *Ras* genes and their protein products

The protein products of the *ras* genes are the prototype for the ras-related superfamily of proteins that share structural homology, this superfamily comprises over 60 mammalian members that have been divided into several subfamilies (ras, rho, rab, arf, ran and rad/gem), division that is based on sequence homology and biological activity. A discussion of these family members is not within the scope of this thesis but is reviewed in (Bourne *et al.*, 1990; Bos 1997).

The ras family consists of the ras proteins (H-ras, *K-ras*A, *K-ras* B and *N-ras*), four rap proteins (rap1A, rap1b, rap2a and rap2b), 3 ras-like proteins (R-ras, TC21 and R-ras3), two ral proteins (ralA and ralB) and the newly identified rheb and M-ras proteins (Bos 1997; Kimmelman *et al.*, 1997; Matsumoto *et al.*, 1997). The ras family is characterised by high similarity in the effector domain and the switch 1 region and all members exhibit GTPase activity (discussed in detail below). The cellular function of this family of proteins is diverse. This chapter will focus on the mammalian *ras* genes; *H-ras*, *K-ras* and *N-ras*.

The *H-ras*, *K-ras* and *N-ras* genes encode proteins of very similar structure and function and are highly conserved between species (Barbacid 1987). These genes are composed of 4 coding exons and 5 non-coding exons (introns); introns differ widely in size and sequence between family members. The human genomic sequences span 3 kb (*H-RAS*), 7 kb (*N-RAS*) and more than 35 kb (*K-RAS*). Two isoforms of K-RAS exist as a result of exon 4 being alternatively spliced (K-RAS 4A and K-RAS 4B), these isoforms diverge at the COOH-terminal 25 amino acid sequence (Barbacid, 1987; Lowy and Willumsen 1993).

The protein product of each of these *ras* genes is approximately 21 kDa and consists of 188 (H-RAS, K-RASA and N-RAS) or 189 (K-RASB) amino acids. These genes are ubiquitously expressed. The level of expression of each gene has been shown to vary in some murine adult tissues and stages of embryonic development and *K-Ras*A exhibits a more restricted pattern of expression suggesting functional differences exist between different Ras proteins (Leon *et al.*, 1987; Furth *et al.*, 1987; Pells *et al.*, 1997). Further evidence in support of functional differences include 1) different *ras* genes are activated in different human tumours suggesting cell-specific activities (Bos 1989; Rodenhuis 1992; discussed in section 1.3.1.1g) 2) regulation of ras activity and transduction of signals to signalling cascades has been noted to be modulated by different proteins i.e. some GTPase proteins and guanine exchange factors show varying affinities between ras proteins, and some effectors of ras exhibit higher affinity for specific ras proteins (Bollag and McCormick 1991; Mizuno *et al.*, 1991; Jones and Jackson 1998; Hamilton and

Wolfman 1998), 3) H-and N-*Ras* knockout mice are completely viable whist *K-Ras* knock out mice die between day 12 and 14 of gestation (Johnson *et al.*, 1997).

1.3.1.1.b Ras proteins are GDP/GTP regulated switches

Ras proteins function as guanosine diphosphate / guanosine triphosphate (GDP/GTP) regulated switches that transduce extracellular ligand mediated stimuli (growth factors, cytokines, hormones and neurotransmitters) to cytoplasmic signal transduction cascades. Ras proteins cycle between two conformations induced by the binding of either GDP or GTP. In the GTP bound conformation, ras is active binding to and activating effector proteins (see below). Ras proteins have low intrinsic hydrolytic activity limiting the rate of GDP/GTP exchange hence cellular control of GDP/GTP cycling is modulated by guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs). GEFs induce the dissociation of guanine nucleotides, and due to the relatively high concentration of cellular GTP to GDP (10:1), this reaction favours, at least initially, the active form of ras-GTP. In turn, ras-GTP is hydrolysed to GDP by intrinsic GTPase activity in combination with GAPs. To date, identified GEFs for ras proteins include sos1, sos2, cdc25, ras-GRF 2 and ras-GRP (reviewed in Malumbres and Pellicer 1998). At least six different GAPs have been identified in mammals: p120-GAP, the neurofibromatosis type 1 protein NF1, GAP1m, GAP11p4BP, IOGAP1 and SynGAP (reviewed in Malumbies and Pellicer 1998).

Analysis of the 3-dimensional protein structure of ras-GDP and ras-GTP in conjunction with the activity of ras proteins that harbour oncogenic mutations has facilitated an understanding as to how GDP/GTP cycling activates the protein. Amino acids 1-164 encode the highly conserved catalytic domain, encompassing the guanine binding site (constituted from amino acids 12-18, 57-63, 116-119 and 114-147) and two "switch" regions (amino acids 30-38 and 60-76). Conformational changes dependent upon GDP or GTP binding are within hydrophilic sites that are located on the external surface of the molecule (encoded by residues 30-38 and 60-76), these regions have been termed switch I and switch II respectively.

Effectors, GAPs and GEFs interact with ras through switch 1 and switch 2. Switch 1 region is the main binding site of ras effector proteins, whilst the switch 2 region has been shown to be responsible for molecular recognition and the specificity of the interaction (Moodie *et al.*, 1995; Nassar *et al.*, 1996; Macara *et al.*, 1996; Akasaka *et al.*, 1996). GAPs and GEFs also interact with the switch regions (Crèchet *et al.*, 1996; Boriack-Sjodin *et al.*, 1998). *In vitro* studies have demonstrated that GAPs and GEFs compete with effector proteins for binding to GTP-ras and that effector protein binding inhibits the activity of GAPs and GTP nucleotide exchange (Marshall, 1996; Giglione and Parmeggiani 1998).

Mutations in *ras*, which result in an oncogenic phenotype, encode proteins that remain in the active GTP-bound state (Scheele *et al.*, 1995). Impaired intrinsic GTPase activity and resistance to GAP activity is associated with mutations at positions 12, 13, 59, 61 whilst mutations at positions 16, 17, 116, 117, 119, 144 and 146 result in increased guanine nucleotide exchange (Bollag and McCormick 1991; Lowy and Willumsen 1993).

1.3.1.1.c Cellular signalling through ras

A large variety of signals, such as growth factors, cytokines, hormones and neurotransmitters, activate ras. In most cases, this activation occurs through cell surface receptors with intrinsic or associated tyrosine kinase activity (Pronk and Bos 1994). In the case of receptors with intrinsic tyrosine kinase activity (RTK), ligand binding stimulates receptor homo- or hetero- dimerisation and/or a conformational change that typically result in autophosphorylation of multiple tyrosine residues in the cytoplasmic portion of the receptor (Pawson 1995). The resulting phosphotyrosine (PY) residues act as highly selective binding sites for so-called 'SH2' (Src homology domain 2) containing proteins, which include the ras specific GEF protein: sos and adaptor proteins GRB2 (growth factor receptor bound protein 2) and SHC (Src homology and collagen) (Pawson 1995; Schenk and Jagalska 1999). The PY residues form a docking site for GEF proteins directly or through binding of adaptor proteins. Protein interactions result in localisation of the GEF to the cell membrane and leads to the transient elevation of ras-GTP levels

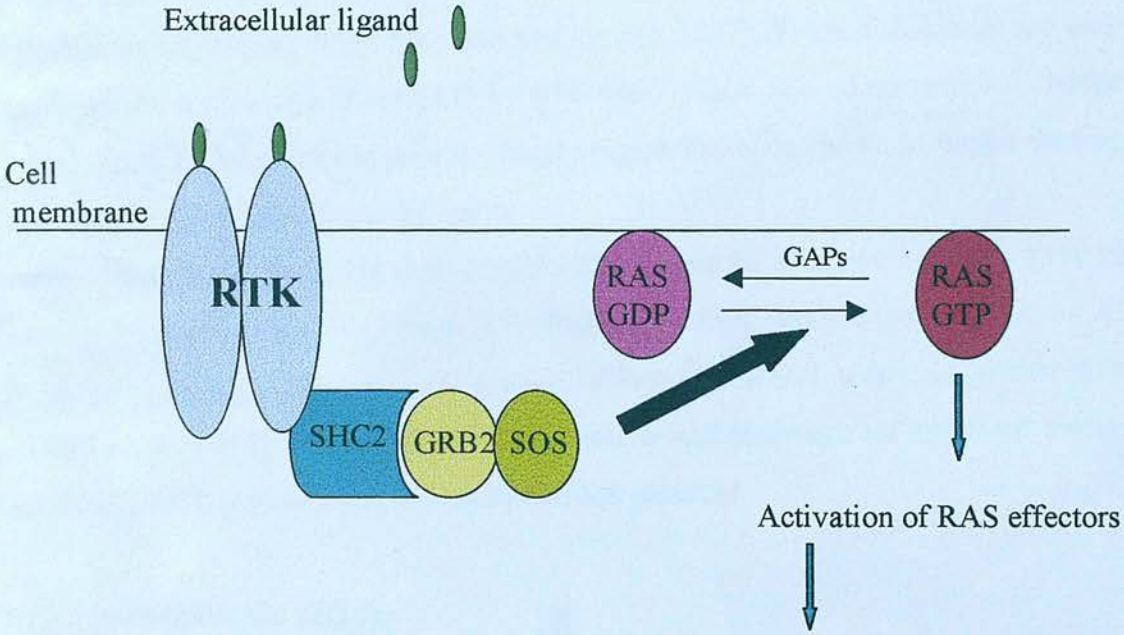
(Schlessinger 1993). Activation of ras through RTK is shown in figure 4.

Receptors not directly associated with tyrosine kinases, such as serpentine receptors are thought to activate ras indirectly through src-like kinases or ligand-independent activation of receptor tyrosine kinases (van Biesen *et al.*, 1995; Lopez-Ilasaca *et al.*, 1997). Calcium and diacylglyceral can activate ras directly by activation of specialised ras-specific GEFs (Farnsworth *et al.*, 1995; Ebinu *et al.*, 1998).

Figure 4: Activation of the ras protein by receptor tyrosine kinases (RTKs)

Binding of ligand to a receptor with tyrosine kinase activity results in receptor dimerisation and autophosphorylation of tyrosine residues in the cytoplasmic portion of the receptor. Phosphorylation of tyrosine residues creates binding sites for SHC, which in turn becomes autophosphorylated and creates recognition sites for the SH2 domain of the adaptor protein GRB2. GRB2 is stably associated with the guanine exchange factor sos. These protein interactions result in the guanine exchange factor, sos, being localised to the plasma membrane, which leads to transient elevation of ras-GTP levels. When in the GTP-bound form, ras protein is active and forms a conformation-allowing interaction with effector proteins that relay the active ras signal downstream through cascades of cytoplasmic proteins.

Figure 4: Activation of the RAS protein by receptor tyrosine kinases (RTKs)



Activation of multiple signalling pathways regulating:

Transcription

Apoptosis

Cytoskeleton

Translation

Cell cycle progression

Golgi trafficking
vesicle formation

1.3.1.1.d Effectors of ras

Numerous proteins have been identified as putative ras effectors. These interact with the GTP-bound form of ras through the "effector loop" (amino acids 32-40), flanking residues and the switch II region of ras protein (Fujita-Yoshigaki *et al.*, 1995; Wittinghoefer and Nassar 1996; Katz and McCormick 1997). These interactions are only possible after the exchange of ras-GDP by GTP which result in a conformational change in switch I and II. Mutations in the ras effector region has been shown to impair binding of the effector proteins as discussed below.

Through different effectors, ras signalling has been found to involve a complex array of signalling pathways involving cross-talk, feedback loops, branch points and multi-component signalling, controlling cell growth, differentiation and apoptosis (reviewed in Campbell *et al.*, 1998). Two relatively well-characterised pathways are discussed below, the raf/MEK/ERK pathway and the ras/PI3K/PKB pathway.

1.3.1.1.e Raf/MEK/ERK pathway

The first effector to be identified in mammalian cells was the protein kinase raf-1, subsequently other family members have been identified (B-raf and A-raf). It is unclear by what mechanism raf effectors are translocated to the nucleus, or how following binding to GTP-ras, raf proteins become activated, although evidence suggests that activation involves a conformational change and modification by phosphorylation, both of which appear to be ras-independent (Stokoe *et al.*, 1994; Morrison and Cutler 1997). Activation of the raf kinases leads to the activation of the MEK/ERK pathway. Following activation, raf phosphorylates and activates mitogen activated protein (MAP)-extracellular signal-related kinase (MEK), which in turn activates the two further extracellular signal-regulated kinases (ERKs), designated $p42^{MAPK}$/ERK2 and $p44^{MAPK}$/ERK1. Once activated ERK kinases translocate to the nucleus and phosphorylate and activate a number of nuclear transcription factors including c-myc which in turn modulate a wide spectrum of

cell processes (Marais *et al.*, 1993; Derijard *et al.*, 1994),

1.3.1.1. f Ras/PI3K/PKB pathway

Another established effector of ras is phosphatidylinositol 3-kinase (PI3K). PI3K is a heterodimer consisting of an 85 kDa regulatory subunit and a 110 kDa catalytic subunit that phosphorylates lipids (Kapeller and Cantley 1994). Interaction with ras-GTP is mediated through the binding of the catalytic subunit of PI3K and the effector region of ras (switch 1 region). Binding of PI3K activates the catalytic activity of the protein and also brings the active protein in close proximity to the plasma membrane. In this location it catalyses the phosphorylation of membrane bound phosphoinositides at the D-3 hydroxyl of myoinositol, producing the second messengers phosphatidylinositol 3-phosphate (PIP), phosphatidylinositol (4,5) (PIP2) and phosphatidylinositol (3,4,5) triphosphate (PIP3). PI3K and the lipid products PIP3 and PIP2 act on pathways that control cell proliferation, cell survival and metabolic changes. PIP2 and PIP3 have several intracellular targets including the protein kinases, protein kinase B (PKB also known as AKT) (Burgering and Coffer 1995; Franke *et al.*, 1995). Experiments have associated PKB kinase activity to the regulation of components of the apoptotic pathway, namely bad (reviewed in Coffer *et al.*, 1998), Caspase-9 (Cardone *et al.*, 1998) and FKHRL1, a member of the Forkhead transcription factor family (Brunet *et al.*, 1999).

A further substrate of PKB, which is of particular interest to this thesis, is glycogen synthase kinase 3β (GSK-3β). Components of the wnt signalling pathway; proteins β-catenin, adenomatous polyposis coli (APC), and Axin and one of the recently identified target genes of this pathway, cyclin D1 are all substrates for GSK-3β. Phosphorylation of β-catenin, APC and Axin mediate β-catenin degradation whilst phosphorylation of cyclin D1 also mediates it's degradation (discussed in detail in section 2.6). *In vitro* experiments have shown that phosphorylation of GSK-3β by PKB results in the inactivation of GSK-3β kinase activity (Cross *et al.*, 1995; Shaw *et al.*, 1997). One may hypothesise that the oncogenic activation of *K-ras* can result in overexpression of PKB, which in turn may

lead to inactivation of GSK-3β kinase activity. Inactivation of GSK-3β kinase activity could disrupt the degradation of β-catenin resulting in increased transcription of target genes (in mammals: *matrilysin*, *cyclin D1* and *c-MYC*) and in the case of cyclin D1 result in increased protein stability. Aberrant expressions of these genes have been detected in NSCLC and are discussed fully in following sections.

1.3.1.1.g Mutations in *ras*

Ras can acquire transforming potential following point mutations that lead to amino acid changes at codons 12, 13 or 61 or through mutations at or near the GTP-binding domain, both of which result in enhanced ability to retain GTP leading to continuous ras signalling (Fayed and O'Brien, 1995). Mutations in the *ras* genes have been detected in a wide variety of tumours, the majority of which occur in the *K-ras* gene.

RAS mutations are observed in approximately one third of cases of NSCLC (Rodenhuis *et al.*, 1987; 1988 Suzuki *et al.*, 1990). In the majority of these cases mutations are in the *K-RAS* gene (>90%) whilst mutations in H-*RAS* and N-*RAS* are relatively rare (Rodenhuis *et al.*, 1988; Suzuki *et al.*, 1990). The frequency of *RAS* mutations reported for NSCLC vary significantly and this may be a result of the relative proportion of different NSCLC subtypes within the cohort. The majority of publications report that *K-RAS* mutations occur primarily in the ADC histological subtype, the reported frequency varies between 20-50% of tumours, whilst in SCC mutations are infrequent (0-13% of cases) (Rodenhuis *et al.*, 1988; Slebos *et al.*, 1989; Reynolds *et al.*, 1991, Mitsudomi *et al.*, 1991; Slebos *et al.*, 1991; Mills *et al.*, 1995; Keohauana *et al.*, 1996; Graziono *et al.*, 1999; Gealy *et al.*, 1999). Conversely, some studies have reported that mutations occur in SCC with a similar frequency to ADC (Rosell *et al.*, 1993; 1996; Gao *et al.*, 1997). It is currently unclear why SCC are mutated in some cohorts but not others. Studies have analysed a full spectrum of SCC, different tumour stages and SCC obtained from smokers and non-smokers. One may suggest that variation in frequency may be a reflection of geographical differences.

The majority of mutations in the *K-RAS* gene occur at codon 12 and the mutation spectrum of this codon consisting of G-T transversions at position 1 (GGT-TGT) in 60% of cases, G-A transition at position 2 (GTT-GAT) in 20% cases and G-T transversion at position 2 (GGT-GTT) in 15% of cases (Rodenhuis *et al.*, 1987; Rodenhuis *et al.*, 1988; Reynolds *et al.*, 1992; Slebos and Rodenhuis, 1992). The majority of mutations detected at codons 13 and 61 are also G-T transversions. The mutation spectrum is consistent with the formation of DNA adducts at guanine residues that can result from the carcinogenic properties of tobacco smoke (see section 1.1.2). In support of this, G-T transversions have been shown to occur more frequently in ADC of smokers than non-smokers (Reynolds *et al.*, 1991), and this mutation is commonly seen in pulmonary ADC of mice treated with constituents of tobacco smoke (You *et al.*, 1989; Belinsky *et al.*, 1997; Nesnow *et al.*, 1998).

It is currently unclear whether mutations in *K-RAS* are an early event in lung carcinogenesis. Studies have shown a homogenous distribution of *RAS* mutation in early stage ADC, suggesting that the mutation is an early event occurring prior to clonal expansion of the tumour (Li *et al.*,1994a). Conversely, studies by Sugio and Co-workers (1994) suggest that *K-RAS* mutations are a relatively late event as gene mutations were detected infrequently in pre-neoplastic lesions associated with SCC; hyperplasia, metaplasia and dysplasia of the bronchioles and central airways. The experimental work carried out for this thesis aims to ascertain whether in our cohort of NSCLC cases there is a statistical significant difference between the frequency of *K-RAS* mutations occurring in ADC of parenchymal or bronchial origin. Furthermore, although analysis of preneoplastic lesions of SCC suggest that mutations are a late event in the development of SCC, this study will analyse AAH lesions, possible precursor lesions of ADC, for mutations within the *K-RAS* gene.

1.3.1.2 MYC

1.3.1.2.a The *C-MYC gene:* structure and function

The *myc* family includes three evolutionary conserved genes, c-, N- and L-*myc,* that have oncogenic potential. These family members are differentially expressed, with *c-myc* being most widely expressed in the majority of pre- and post-natal proliferating cells (reviewed in Marcu *et al.*, 1992). Expression is strictly dependent upon mitogenic stimuli and is down regulated in many cell types when induced to terminally differentiate (Evan and Littlewood, 1993; Ryan and Birnie, 1997). Oncogenic activation is generally a result of their constitutive expression, and overexpression is noted in a large fraction of human malignancies (reviewed in Nesbit *et al.*, 1999). Investigations into the cellular role of myc proteins have focused on the *c-myc* gene. Studies have demonstrated that deregulation of *c-myc*, rather than just overexpression is sufficient for oncogenic activation through the disruption of its role in the control of cellular proliferation, apoptosis and differentiation (Reviewed in Sakamuro and Prendergast 1999).

The protein product encoded by the *c-myc* gene is a transcription factor. The protein can be divided into 3 regions; the N-terminal domain (NTD), central domain and carboxyl terminal domain (CTD), both the NTD and CTD are essential for transcriptional transactivation, proliferation and transformation whereas the central portion of the protein has been shown to be dispensable (Henriksson and Luscher 1996).

The carboxyl terminus domain (CTD) contains a leucine zipper (Zip), a helix-loop-helix (HLH) motif and an adjacent domain rich in basic amino acids (Murre *et al.*, 1989; Prendergast and Ziff, 1989; Blackwell *et al.*, 1990). These motifs are contiguous and are referred to as the bHLHZip region (Henriksson and Luscher 1996). The HLHZip motifs mediate protein oligomerisation. Myc proteins do not homodimerise *in vivo,* but oligermerise with a number of other proteins that also contain HLHZip motifs. The first protein identified was max, a ubiquitously expressed protein that is comprised essentially of a bHLHZip domain (Blackwood *et al.*, 1991). Heterodimerisation with max is necessary for c-myc to mediate gene transactivation, transformation, cell cycle

progression and apoptosis (Amati *et al.*, 1993; Packham and Cleveland 1995). Included among the target genes transactivated by myc/max are cdc25a, *cyclin A*, *cyclin E*, *α-prothymosin, lactate dehydrogenase A* and *orthinine decarboxylase* (*ODC*) (reviewed in Grandori and Eisenman, 1997; Hoffman and Liebermann, 1998).

In addition to max, several other proteins, including the tumour suppressor gene BRCA1, have been implicated in interactions with this region. All these proteins have been linked to transcriptional regulation (reviewed in Sakamuro and Prendergast, 1999). Unlike interaction with max, interaction of myc with these proteins results in inhibition of c-myc transactivation function.

The N-terminus domain (NTD) acts as the transactivation domain. This NTD contains multiple phosphorylation sites and myc-homology Box 1 (MB1) and Box 2 (MB2) domains that are implicated in the regulation of c-myc activity (Lutterbach and Hann 1994; 1997; reviewed in Sakamuro and Prendergast, 1999). The NTD has been shown to interact with a number of proteins *in vivo*, to date seven proteins have been identified (reviewed in Sakamuro and Prendergast, 1999). The interaction of myc with these regulatory and effector proteins link c-myc to numerous cellular processes such as the cell cycle, chromatin modelling, apoptosis or cell fate signalling.

1.3.1.2.b *Myc* genes in NSCLC

Myc genes have been shown to be activated by a number genetic alterations including chromosomal translocations, proviral insertion, retroviral transduction and gene amplification (Ryan and Birnie, 1997). *N-MYC* and *L-MYC* do not appear to be deregulated in NSCLC but *c-MYC,* which is located at chromosome region 8q24, has been shown to be overexpressed in 50% of cases (Cline *et al.*, 1987; Slebos *et al.*, 1989; Paget 1989; Field and Spandidos 1990; Prins *et al.*, 1993). The *c-MYC* gene has been shown to be amplified in 10% of these cases but in the remaining cases the mechanism of gene deregulation is unknown. *C-MYC* is a target gene of the MEK/ERK pathway (see section 1.3.1.1.e), therefore gene expression may be deregulated due to activating mutations in *ras* or by deregulation of any other components of this pathway.

C-myc has recently been shown to be a target gene of the wnt signalling pathway (He *et al.*, 1998). This discovery offered an explanation as to why *c-MYC* was overexpressed in approximately 70% of colorectal cancer cases, and as with NSCLC, the frequency of gene amplification was low and no other genetic alterations had been detected (Erisman *et al.*, 1985; Rothberg *et al.*, 1985; Sikora *et al.*, 1987). The majority of colorectal tumours contain mutations in either of two key components of this pathway, *APC* or *β-catenin*, mutations that result in increased expression of target genes, discussed in detail in section 2.6 (Sparks *et al.*, 1998). **To my knowledge there have been no reports regarding mutations in the wnt signalling pathway in NSCLC which may result in overexpression of *c-MYC*. However, loss of heterozygosity has been detected at polymorphic sites within the *APC* gene (discussed in section 1.3.3.3) and experimental work included in this thesis aims to determine whether the *APC* tumour suppressor gene has a role in NSCLC.**

1.3.1.3. Protein Kinases

NSCLC cases have been shown to overexpress the protein kinases *c-ERB-B1*, *c-ERB-B2* and *c-MET*.

1.3.1.3a c-*ERB* family

Erb-b1 and erb-b2 are members of the erb-b family of transmembrane receptor tyrosine kinases (Alroy *et al.*, 1997). Signalling by these receptors, and two other family members (*erb-b3* and erb-b4), can be stimulated by at least six different epidermal growth factor (EGF) ligands. Binding of extracellular ligand result in conformational alteration of the extracellular domain which induces receptor oligomerisation. Oligomerisation stabilises interactions between adjacent cytoplasmic domains and leads to elevated protein kinase activity and enhanced ligand-binding activity (Yarden and Schlessinger, 1987; Schlessinger, 1988). Following ligand binding, specific tyrosine residues within the carboxyl tail of the receptor are autophosphorylated and proteins containing SH2 domains are recruited to the activated receptor. Specific SH2 proteins may be recruited depending upon which receptor family member is activated, for example autophosphorylation of erb-b3 recruits PI3K thus erb-b3 heterodimers will selectively activate the PI3K pathway (Carraway *et al.*, 1995). Numerous signalling pathways can be activated in this way including pathways mediated by ras (see section 1.3.1.1).

Erb-b1 encodes the epidermal growth factor receptor (EGFR). This receptor has been shown to be overexpressed in approximately 45% of NSCLC, gene amplification is uncommon and the mechanism that results in deregulated expression is unclear (Kern *et al.*, 1992; Volm *et al.*, 1992; Rusch *et al.*, 1993; 1995). Overexpression of the *EGFR* ligand, TGFα, has also been reported to occur in approximately 10% of NSCLC cases that do not aberrantly express *EGFR*, whilst overexpression of both *EGFR* and *TGFα* have been detected in 38% of cases. Cell type and stage do not influence differential expression of *EGFR* or *TGFα*, and studies have shown overexpression of *EGFR* in a

spectrum of preneoplastic lesions including metaplasia, dysplasia and carcinoma *in situ* (Rusch *et al.*, 1995; 1997). Therefore, it appears that deregulation cellular signalling through EGFR is an early event in NSCLC.

The transmembrane tyrosine kinase receptor encoded by *erb-b2* (*her2/neu*) does not have a recognised ligand but has been shown to dimerise with other members of the erbb family and may be activated autonomously if receptor levels are high (Tzahar *et al.*, 1994; Rachwal *et al.*, 1995). Overexpression of *ERB-B2* has been detected in NSCLC, with some studies suggesting a higher percentage of overexpression in ADC than in SCC (Kern *et al.*, 1990; Tateishi *et al* 1991; Shi *et al.*, 1992). *ERB-B2* overexpression may be an early event in the development of ADC as in AAH lesions aberrant expression has been detected in 7-35% of cases, this expression has been correlated with increase atypia (Kerr *et al.*, 1994; Kitamura *et al.*, 1996).

1.3.1.2.b c-Met

The *c-met* gene encodes a membrane receptor tyrosine kinase that is activated by hepatocyte growth factor (HGF), a cytokine that is mitogenic and a motogenic factor for both cultured bronchial epithelial cells and non-small-cell carcinoma lines (Brinkmann *et al.*, 1995; Singh-kaw *et al.*, 1995). Overexpression of the receptor is observed in approximately 50% of NSCLC, of which overexpression is most frequent in ADC (30%) and LCC (20%) and is rare in SCC (Prat *et al.*, 1991; Liu and Tsao 1993; Tsao 1998). HGF is also overexpressed in these NSCLC subtypes (Olivero *et al.*, 1996). Among ADC, intermediate to high levels of c-MET immunoreactivity correlated with a greater degree of tumour differentiation. Furthermore, an accentuation of c-MET immunoreactivity was often noted in cancer cells at the advancing edge of tumours. These findings support a role for the c-MET tyrosine kinase receptor in lung cancer cell invasion and differentiation *in vivo* (Tsao *et al.*, 1998).

1.3.1.4 *BCL2* and apoptosis

BCL2 was originally identified at the chromosomal breakpoint of t(14;18) in a lymphocytic B cell leukaemia (Tsujimoto *et al.*, 1984; Cleary *et al.*, 1986). Since its identification *BCL2* has been shown to be one of a family of genes, of which there is at least 16 members in humans, that is involved in the regulation of apoptosis. Some family members enhance (e.g. *BAX* and *BAD*) and some negatively regulate apoptosis (e.g. *BCL2* and *BCLXL*), reviewed in Kroemer, 1997. Bcl2 family members can form hetero- and homo- dimers and it has been postulated that the balance of pro-verses anti-apoptotic family members dictates cellular response. The current understanding as to how the bcl2 family regulate apoptosis is reviewed in Reed, 1998.

Several studies have analysed the expression of *BCL2* and found that protein overexpression is significantly associated with SCC (35-59%) and is less frequent in ADC (10-28%) and LCC (38%) (Jiang *et al.*, 1996; Higashiyama *et al.*, 1997; Apolinario *et al.*, 1997; Laudanski., 1999; Dosaka-Akita., 1999). Aberrant expression of *BCL2* is thought to be an early event in SCC as overexpression of the protein has been detected in dysplastic lesions (Walker *et al.*, 1995).

1.3.1.5 Cyclin D1

The *CCND1* proto-oncogene, located on chromosome 11q13, encodes cyclin D1, a protein that is a major regulator of cell proliferation (reviewed in Sherr 1996 and discussed in sections 11.3.2.1 and 1.3.2.4). Following mitogen stimulation, cyclin D1 and other D-type cyclins are induced and assemble with their catalytic partners, cyclin dependent kinases, CDK4 and CDK6. These holoenzyme complexes enter the nucleus where following phosphorylation by a CDK-activating kinase (CAK), function to phosphorylate protein substrates. The major targets of the cyclin D-CDK complexes are the retinoblastoma family of proteins RB, p107, and p130 (see section 1.3.2.1). The most recognised function of cyclin-D dependent kinases is phosphorylation of the protein

product pRB that leads to further phosphorylation by other cyclin dependent kinases and results in hyperphosphorylation of the protein. Whilst hyperphosphorylated, the growth suppressive functions of RB are disrupted allowing transcription of a number of genes whose activities are required for DNA synthesis (reviewed in Dyson 1998; Nevins 1998).

In vitro experiments implicate *CCND1* as a proto-oncogene and studies have shown that cyclin D1 overexpression confers tumourigenicity (Musgrove *et al.*, 1994). Elevated levels of cyclin D1 have been reported to occur in many types of human cancer including cancers of the breast, head and neck, colon and lung (Motokura and Arnold 1993; Hunter and Pines 1994; Hall and Peters, 1996). In some cases, overexpression of the gene has been shown to be a result of chromosomal translocations or gene amplification (Schuuring *et al.*, 1992; Buckley *et al.*, 1993; Zhang *et al.*, 1993). Amplification of *CCND1* has been detected in 5-30% of NSCLC cases whilst overexpression of cyclin D1 protein occurs in approximately 50% of cases (Betticher *et al.*, 1996; Marchetti *et al*; 1998; Reissmann *et al.*, 1999; Mishina *et al.*, 1999). Overexpression of cyclin D1 in cases of NSCLC that do not contain amplification of the *CCND1* gene, may be a result of loss of *RB* expression (see section 1.3.2.1), disruption of wnt (see section 2.6.10.6) or ras signalling pathways (see section 1.3.1.1.f).

No significant difference has been demonstrated between cyclin D1 overexpression and histological subtype, and aberrant expression has been detected at all tumour stages suggesting cyclin D1 overexpression is an early event in NSCLC.

1.3.2 Tumour suppressor genes and lung cancer

Tumour suppressor genes are cellular genes that inhibit or restrain the process of tumourigenesis. Knudson proposed the concept of tumour suppressor genes over 20 years ago (Knudson 1978). By studying the inheritance patterns of retinoblastoma he concluded an inherited mutation in one allele of the tumour suppressor gene *RB*, predisposed the individual to the development of retinoblastoma. This mutation was recessive as development of the disease required the subsequent inactivation of the remaining allele i.e. by a somatic event resulting in deletion or mutation of the remaining 'active' allele. In sporadic cancers two somatic events resulting in the inactivation of both alleles of a tumour suppressor gene are required, in most cases of sporadic cancer one allele is deleted whilst the other is inactivated through DNA mutation (Weinberg, 1991). This section reviews the tumour suppressor genes implicated in NSCLC.

1.3.2.1 The retinoblastoma gene

The retinoblastoma (*RB*) tumour suppressor gene, located on chromosome 13q14.11, was one of the first tumour suppressor genes identified (Knudson, 1978; Lee *et al.*, 1987). Individuals that harbour germline mutation within the gene are predisposed to the development of tumours of the retina (retinoblastoma), a small proportion of these individuals also develop osteosarcomas. The *RB* gene encodes a 105 kDa nuclear phosphoprotein (pRB) (Yunis and Ramsay, 1978). The protein functions to inhibit the cell cycle through the control of progression through late G1 phase. pRB physically interacts with, and inhibits the action of cellular proteins, such as the E2F family of transcription factors, which function to activate genes required for DNA synthesis during the S phase of the cell cycle (Strauss *et al.*, 1995; Bartek *et al.*, 1996; Ikeda *et al.*, 1996; Wang 1997; Herwig and Strauss 1997). The activity of pRB as a repressor of G1 progression is regulated by cycles of phosphorylation and dephosphorylation at multiple serine and threonine residues within the RB protein (Bartek *et al.*, 1996; Wang 1997).

During G0/G1 phase of the cell cycle, pRB is hypophosphorylated and becomes phosphorylated by the cyclin D-dependent kinases, CDK4 and CDK6, as cells approach the G1/S boundary prior to DNA synthesis. Phosphorylation of pRB leads to a release of active forms of the E2F family of transcription factors, which can in turn activate target genes such as cyclin E, cyclin A and many S phase-specific genes such as thymidine kinase and polymerase α (reviewed in Lavia and Jansen-Durr 1999).

Mutations in the *RB* gene have been detected in a wide variety of sporadic cancers including lung carcinoma. In SCLC and NSCLC the inactivation of both *RB* alleles is well documented with loss of *RB* expression reported to occur in 17-41% of NSCLC tumours (Xu *et al.*, 1991; 1994; Reissmann *et al.*, 1993; Sakaguchi *et al.*, 1996; Gorgoulis *et al.*, 1998; Tanka *et al.*, 1998; Vonlanthen *et al.*, 1998). Aberrant expression of the *RB* gene has been correlated with NSCLC stage, with 20% of stage I and II tumours and 60% of stage III and IV showing loss of expression (Xu *et al.*, 1991). This suggests that *RB* may be important in cancer progression rather than initiation, a proposal that is supported by the fact that individuals heterozygous for mutations in *RB* do not have an increased risk of lung cancer.

The RB protein is also thought to act as a positive transcriptional regulator of the cyclin D1 gene, *CCND1*, in some cell types (Muller *et al.*, 1994; Lukas *et al.*, 1995). Analysis of NSCLC has failed to correlate the presence of pRB with varying levels of cyclin D1 detected in these tumours (Schauer *et al.*, 1994). One may hypothesise that in NSCLC and possibly other tumour types the RB and cyclin D1 levels do not correlate due to deregulation of other pathways that influence the protein levels of these two genes, i.e. in the case of increased levels of cyclin D1 one could specify pathways such as MEK/ERK (see section 1.3.1.1.e), PI3K (see section 1.3.1.1.f) or wnt (see section 2.6.10.6).

1.3.2.2 *p53*

The *p53* gene, located at chromosome 17p13.1, encodes a 53 kDa nuclear phosphoprotein which functions as a transcription factor that is involved in the regulation of a variety of cellular functions including DNA replication, transcription, cell cycle

progression, differentiation and apoptosis (Reviewed in Kirsch and Kastan 1998; Steele *et al.*, 1998; Almog and Rotter 1998). The *p53* gene has been termed "guardian of the genome" as following DNA damage the cell rapidly increases its level of p53 by a post-transcriptional mechanism which results in induction of cellular genes that induce G1 arrest or programmed cell death (Lane 1992; Levine 1997). Further to this, inactivation of p53 has been shown to result in increased frequencies of gene mutations, chromosomal rearrangements and abnormal chromosomal segregation (Havre *et al.*, 1995; Fukasawa *et al.*, 1996; Bertrand *et al.*, 1997).

Germline mutations in the *p53* gene are found in patients with the autosomal dominant Li-Fraumeni cancer syndrome, whereby affected individuals have early onset of multiple tumours (Li *et al* 1988; Malkin *et al.*, 1990; 1993; 1994). Somatic mutations have been detected in various cancers including pulmonary cancer (Takahashi *et al.*, 1989; Hollstein *et al.*, 1991). *P53* mutations are of two types: dominant negative mutations that stabilise the protein and recessive null mutations. Usually within the cell the level of p53 is very low and difficult to detect by immunological techniques. In cells with dominant negative mutations that result in increased half-life of the protein, immunological techniques are able to detect the protein.

Mutations within the *p53* gene have been shown to occur in approximately 48% of NSCLC, with a significant difference between the frequency of mutations in SCC (67%) and ADC (39%) (Greenblatt *et al.*, 1994, reviewed in Fujita *et al.*, 1999). Most mutations, as with cancers at other sites, are located between codons 120 and 290, exons 5 to 8 of the gene (Chiba *et al.*, 1990; Greenblatt *et al.*, 1994). This region of the gene encodes the central core domain and mutations in this region have been shown to lead to the loss of p53 DNA binding activity (Kern *et al.*, 1992; Velculescu and El-Deiry, 1996). Within this region several mutational hotspots have been detected, these hotspots occur at codons 175, 248, 273 and 281/282 (Greenblatt *et al.*, 1994). Mutations detected in NSCLC are commonly G-T transversions (Greenblatt *et al.*, 1994; Huang *et al.*, 1998; see section 1.1.2).

Immunohistochemistry has been used to determine the timing of *p53* alterations within

SCC tumourigenesis. Numerous studies have detected p53 protein in dysplastic lesions and the frequency of p53 detection correlates with increasing dysplasia and is reported to occur in over 70% of CIS, suggesting that aberrant expression of *p53* is important in driving progression of severe bronchial dysplasia to an invasive SCC (Bennett *et al.*, 1993; Nuorva *et al.*, 1993; Hirano *et al.*, 1994). However, these data must be interpreted cautiously as p53 accumulation may be a reflection of initiation of G1 arrest or apoptosis following cellular damage. Confirmation that the *p53* gene is mutated in dysplastic lesions has come from several studies (Sozzi *et al.*, 1992; Chang *et al.*, 1995; Franklin *et al.*, 1997).

Due to the lack of identified precursor lesions, it is unclear whether mutation in the *p53* gene is an early event in ADC formation. Analysis of ADC has suggested that *p53* mutations may occur prior to clonal expansion, as determined by their homogeneous topographical distribution (Li *et al.*, 1994a). Several groups have assessed p53 accumulation in AAH lesions and reports show aberrant expression in 5-58% of lesions (Kerr *et al* 1994; Kitamura *et al.*, 1996, Cagle *et al.*, 1996; Slebos *et al.*, 1998), in several studies aberrant expression correlated with increased atypia. The large variation in reported frequency of aberrant expression could be due to several factors including antibody and immunohistochemical technique used, interpretation of positive staining and the relative percentage of lesions showing high grade and low grade atypia.

1.3.2.3 The *CDKN2* gene and its protein products p16 and $p14^{ARF}$

The *CDKN2* gene, mapping to chromosome 9p21, has been termed a classic tumour suppressor gene. Germline mutations within the *CDKN2* gene are associated with familial multiple-mole melanoma and the resulting tumours exhibit biallelic inactivation (Hussussian *et al.*, 1994). The chromosome region 9p21 has also been shown to be frequently deleted in primary tumours and tumourigenic cell lines (Hussussian *et al.*, 1994; Nobori *et al.*, 1994; Kamb *et al.*, 1994; Testa, 1996). The region of loss identified within these tumours not only encompasses the *CDKN2* gene but also the *MTS 2/ INK4B* (*multiple tumour suppressor gene 2*) (Xiao *et al.*, 1995; Stone *et al.*, 1995a;b).

The *CDKN2* gene is very unusual in that it encodes two distinct proteins translated from alternatively spliced mRNA, each regulated by its own promoter (Harber, 1997). P16 ($p16^{INK4a}$) is specified by a transcript containing exons 1, 2 and 3, and the alternative, or ß transcript containing exons 1ß, 2 and 3. The alternative, or β transcript encodes the protein termed $p19^{ARF}$ in the mouse (Quelle *et al.*, 1995) and $p14^{ARF}$ in humans (Duro *et al.*, 1995; Mao *et al.*, 1995; Stone *et al.*, 1995b). P16 and $p14^{ARF}$ have recently been identified as playing crucial roles in the regulation of the cell cycle; p16 is associated with the *RB* pathway (see section 1.3.2.1) whilst $p14^{ARF}$ is involved in the regulation of *p53* pathway (1.3.2.2).

P16 consists of a 156 amino acid protein that functions as a cyclin dependent kinase inhibitor (CDI). P16 is representative of a class of CDIs with amino acid homology (p16, p15 (encoded by *MTS2*), p18, p19) that bind CDKs, inhibiting their ability to phosphorylate the RB protein and resulting in G1 cell cycle arrest (Hannon and Beach, 1994). Evidence suggests that the RB protein negatively regulates p16 and an inverse correlation between alterations in these proteins has been determined in several tumour studies (Li *et al.*, 1994b; Tam *et al.*, 1994; Shapiro *et al.*, 1995).

CDKN2 also transcribes $p14^{ARF}$ (alternative reading frame) which encodes, in humans, a 15 kDa protein. $P14^{ARF}$ has only recently been described as having tumour suppressor functions. In NIH3T3 fibroblasts overexpression of the murine $p19^{ARF}$ homologue has been shown to induce G1/S and G2/M cell cycle arrest (Quelle *et al.*, 1995) whilst $p19^{ARF}$ knockout mice develop carcinogen-induced tumours at an accelerated rate (Kamijo *et al.*, 1997).

The protein product $p14^{ARF}$ binds specifically to the protein MDM2. MDM2 is involved in the regulation of S phase transition through its ability to bind to the p53 protein preventing p53 mediated induction of genes required for G1 cell cycle arrest i.e. p21. MDM2 when in a protein complex with $p14^{ARF}$ can still associate with p53, however this protein complex ($p53/MDM2/p14^{ARF}$) is more stable that the p53/MDM2 complex, which results in inhibition of G1 cell cycle arrest (Kamijo *et al.*, 1997; Stott *et al.*, 1998).

Early analyses of the *CDKN2* gene comment only on p16 as researchers were unaware of

$p14^{ARF}$; therefore most data to date discuss mutations in relation to p16 and address the aberrant expression of p16 RNA and protein.

P16 expression has been described in the normal lung, expression is nuclear with weaker cytoplasmic staining (Merlo *et al.*, 1995; Kashiwabara *et al.*, 1998; Gazzeri *et al.*, 1998). Aberrant expression, as depicted by failure to detect protein product, has been frequently observed in NSCLC occurring in up to 66% of cases (Kratzke *et al.*, 1996; Sakaguchi *et al.*, 1996; Tanka *et al.*, 1998; Gazzeri *et al.*, 1998; Kashiwabara *et al.*, 1998). The mechanisms by which p16 expression is lost include genomic alteration in exons 1α and 2 (0-14% of cases), homozygous deletions (45% of cases) and methylation of exon 1α. (30-70%) (Stirzaker *et al.*, 1997; Tanka *et al.*, 1998; Gazzeri *et al.*, 1998). Loss of p16 expression is associated more frequently with SCC than ADC (Kinoshita *et al.*, 1996; Kratze *et al.*, 1996; Sakaguchi *et al.*, 1996; Kashiwabara *et al.*, 1998) and has been correlated with degree of differentiation in ADC, with p16 expression more common in poorly differentiated than well-differentiated adenocarcinomas (Kashiwabara *et al.*, 1998).

In the majority of human cell lines, large deletions have been shown to occur that encompass the coding region of p16 and $p14^{ARF}$, although in a few cases, point mutations or small deletions that effect p16 only have been detected (Quelle *et al.*, 1997). This may suggest that loss of both p16 and $p14^{ARF}$ may favour tumour formation (Stone *et al.*, 1995b).

1.3.3 Gross chromosomal abnormalities, fragile sites and loss of heterozygosity

Cytogenetics and molecular biology techniques have demonstrated that NSCLC cases are often aneuploid and exhibit numerical and structural changes in chromosomes. These tumours may be described as exhibiting chromosomal instability (CIN), a phenotype originally described in colorectal cancers. Chromosomal instability is thought to be a result of abnormalities in genes that encode proteins that function in cell cycle check points, e.g. *p53* (Cahill *et al.*, 1998). Numerous abnormalities occurring at specific chromosome regions are associated with NSCLC, some of these regions are known

"fragile sites", i.e. sites that display a number of characteristics of unstable, highly recombinogenic DNA, and have been shown to be preferential sites for chromosomal deletion and rearrangement (Smith *et al.*, 1998). Chromosomal amplifications and deletions associated with NSCLC are discussed below.

1.3.3.1 Regions of DNA amplification

Amplification of proto-oncogenes is a common feature of many solid tumours (Alitalo and Schwab, 1986; Schimke, 1988). Chromosomal regions of DNA amplification can be identified by the presence of double minutes or homogeneously staining regions or through the use of the relatively new technique comparative genomic hybridisation (CGH) analysis (reviewed in Houldsworth and Chaganti 1994). These techniques have been used to analyse primary NSCLC tumours and NSCLC cell lines (Balsara *et al* 1997, Taguchi *et al.*, 1997). A high incidence of over representation of chromosome arms 3q (85%), 5p (70%), 7p (65%) and 8q (65%) were detected in these studies. DNA sequence amplification is most frequent at 3q26, the site of the *PIK3CA* gene that encodes the catalytic subunit of phosphatidylinositol-3 kinase (see section 1.3.1.1.f), whilst chromosomal amplification at 8q24 (which harbours *c-MYC*; see section 1.3.1.2) was detected in 20% of cases. The chromosomal region 7p12, containing *ERB-B1* was amplified in 50% of cases and of these cases, two thirds also exhibited chromosomal amplification at 2p13, a region containing the *TGFα* gene. (See section 1.3.1.3a).

1.3.3.2 Regions of chromosomal loss

Cytogenetic and comparative genomic hybridisation (CGH) studies report allelic loss at many chromosomal regions in NSCLC cases. The use of molecular techniques has been used to narrowly define the critical regions of loss. Initial analysis utilised polymorphic loci based on restriction fragment length polymorphisms (RFLP). These polymorphic sites however were at uncertain sites throughout the chromosome and most were of low informity (i.e. the polymorphic variants occurred at a very low frequency). A further drawback was that RFLP were detected by southern blotting which required large

amounts DNA. Since these initial studies, the number and informativeness of RFLP markers has increased dramatically and information has been further enhanced by new types of markers utilising highly polymorphic DNA sequences, such as variable number of tandem repeat (VNTR) loci (Nakamura *et al.*, 1987), microsatellite repeats containing variable copy numbers of cytosine and adenosine repeats ((CA)n) (Weber and May 1989) and tri or tetranucleotide repeats (Gastier *et al.*, 1995; Dubovsky *et al.*, 1995). These loci have multiple alleles resulting in a frequency of heterozygosity that can be 80% or higher. With the introduction of polymerase chain reaction (PCR) and the development of PCR-based polymorphic markers, it is now possible to detect genetic change even in the smallest number of cells. This is further refined by microdissection, which allows purification of tumour cells from normal stromal cells. Using these refined techniques more than 30 regions of deletion, either loss of heterozygosity (LOH) or homozygous deletions, dispersed on 21 different chromosome arms have been defined in NSCLC (Reviewed in Kohno and Yokota 1999), the majority of these deletions occur within common fragile sites (Smith *et al.*, 1998).

The most common losses identified in NSCLC lie on chromosome 3p, 13q and 17p, frequency of loss has been detected in >70% of cases. Loss of heterozygosity include 6q, 9p and 19p in >60% cases. Tumour suppressor genes have already been identified on 13q (*RB*), 17p (*p53*) and 9p (*CDKN2*), and have been discussed in detail in previous sections. These studies analysed multiple polymorphic sites along the chromosome arm and demonstrated multiple regions of loss; further studies are now needed to elucidate tumour suppressor gene in these regions.

The allelotypes of SCC and ADC have been shown to vary (Sato *et al.*, 1994). Loss of heterozygosity is more frequent in SCC than in ADC as determined by the analysis of 36 chromosomal arms. In SCC most frequent losses are seen at chromosome 17p (88-100%), 3p (82-100%), 9q (67%) and 13q (60%), whilst in ADC loss at 17p (51%), 3p (40%), 9p (36%) and 8q (32%) are most frequent (Tsuchiya *et al.*, 1992; Sato *et al.*, 1994). The frequencies of LOH on chromosomes 17p and 3p are significantly different between SCC and ADC, whilst LOH on chromosome 2q has been reported to only occur

in ADC (Tsuchiya *et al.*, 1992; Sato *et al.*, 1994). These data suggest that different tumourigenic mechanisms may be involved in histological types of NSCLC.

Losses of 3p, 13q and 17p have been associated with stage 1 primary NSCLC. A study specifically addressing early genetic events in SCC analysed normal epithelium, preneoplasia (hyperplasia, metaplasia and dysplasia) and carcinoma *in s*itu (CIS), from smokers, for LOH at 10 chromosomal regions frequently deleted in NSCLC. These chromosomal regions were: 3p12, 3p14.2-21.3, 3p21, 3p22-24, 3p25, 5q22 (region of the *APC* and *MCC* genes, see below), 9q21 (*CDKN2*), 13q14 (*RB*) and 17p13 (*p53*). These studies determined that LOH at 3p21, 3p22-24, 3p25 and 9p21 were the most frequent events and occurred in the earliest detectable lesions, in fact they were also detected in histologically normal epithelium. The incidence of LOH within a sample increased with increasing morphological change i.e. dysplasia and CIS demonstrated a significant increased overall incidence of LOH than normal epithelium or metaplastic lesions, falling in line with the theory that a sequence of molecular events precedes the development of cancer. Losses at other regions of 3p and the 17p13 (*p53* gene) were present mainly in histologically advanced lesions (dysplasia and CIS). In contrast deletions at chromosome 5q22 and 13q (near the *APC* and *RB* genes) were mainly associated with invasive lesions (Wistuba *et al.*, 1999).

1.3.3.3 LOH at 5q21

Several groups have reported genetic loss at chromosome 5q21 in NSCLC (Ashton-Rickardt *et al.*, 1991; D'Amico *et al.*, 1992; Tsuchiya *et al.*, 1992; Horri *et al.*, 1992a; Hosoe *et al.*, 1994; Wieland and Bohm 1994, Fong *et al.*, 1995). These studies report LOH in 20-71% of cases. The varying frequency of LOH may be a reflection of different histological types within cohorts as a number of studies have revealed that ADC and SCC exhibit differences in allelotypes (see section 1.3.3.2). Tumour stage may also be a cause for this variation, as discussed in the previous section the analysis of different stages of SCC has determined that different genetic events occur at different stages of tumour development. This phenomenon has been well described during pathogenesis of colon

cancer where both the accumulation of mutations and the order have been shown to be important in the colon tumourigenesis (Fearon and Vogelstein 1990; Kinzler and Vogelstein 1996).

Two studies have reported a correlation between LOH at 5q21 and ADC and SCC histological types, however the results from these groups were conflicting (Tsuchyiya *et al.*, 1992; Fong *et al.*, 1995). Fong and co-workers also reported that LOH at 5q21 occurs more frequently in late stage tumours, suggesting that loss of this chromosomal region is a late event in the tumourigenesis of NSCLC.

1.3.3.3.a Putative tumour suppressor genes: *APC* and *MCC*

Several genes have been identified that reside within the chromosomal region 5q21. These genes are *FER*, *TB1*, *MCC*, *TB2*, *SRP19* and *APC* (Kinzler *et al.*, 1991a,b; Joslyn *et al.*, 1991). Chromosome region 5q21 was initially investigated to identify the tumour suppressor gene responsible for the inherited disorder, familial adenomatous polyposis coli (FAP). FAP is a disease whereby multiple polyps develop within the colon predisposing the affected individual to colonic carcinoma and increased risk of cancers of the brain, thyroid, bone and focal proliferative lesions of the connective tissue (Bodmer *et al*, 1987; Leppert *et al.*, 1987; Dunlop *et al.*, 1990). FAP kindred's were screened for mutations within the six identified genes, mutated in colorectal cancer (*MCC*), *TB1, SRP19, TB2* and the adenomatous polyposis coli gene (*APC*).

Mutations in the *MCC* gene were detected in sporadic colon tumours (Kinzler *et al*, 1991a) although further larger studies failed to show mutations within this gene (Curtis *et al.*, 1994; Cripps *et al.*, 1995). No mutations were detected in the *TB1, SRP19, TB2* genes (Nigro *et al.*, 1989; Nishisho *et al.*, 1991). Germline mutations were detected in the gene that was subsequently termed the *APC* gene. In the majority of familial and sporadic colorectal cancer cases both *APC* alleles were shown to be inactivated either by gene mutation alone or by loss of one allele and mutation within the remaining allele (Nishisho *et al.*, 1991; Groden *et al.*, 1991). Biallelic inactivation of *APC* has also been reported in sporadic gastric (Tamura *et al.*, 1994) and hepatic cancers (Oda *et al.*, 1996;

Imai *et al.*, 1997), suggesting that inactivation of the *APC* gene may be a critical step in the genesis of several types of tumours. It is unknown whether biallelic inactivation of *APC* is a feature of NSCLC. One study consisting of 7 NSCLC cases, that displayed LOH at 5q21, detected no mutations in the remaining *APC* allele (Horri *et al.*, 1992b). **Therefore, experimental work undertaken for this thesis investigates whether biallelic inactivation of the *APC* gene is a feature of NSCLC.**

Chapter 2: The *APC* gene

This chapter reviews the current knowledge on the expression of *APC* and its protein structure. Several proteins have been identified that interact with APC whilst other interactions are still putative. These protein/protein interactions are discussed and possible cellular functions are addressed. The mutation spectrum identified within the *APC* gene in both familial and sporadic colorectal cancer and in other sporadic cancers is reviewed. Publications have described the identification and characterisation of homologues of *APC*, these homologues are discussed in detail in section 2.5.

2.1 *APC* expression

The *APC* gene was originally identified to encode a 2843 amino acid polypeptide with a predicted molecular weight of 310 kDa (Groden *et al.*, 1991; Kinzler *et al.*, 1991a). Recent investigations have revealed that several isoforms of APC exist (discussed in section 2.2) and protein products of various molecular weights have been identified (Santoro and Groden 1997; Pyles *et al.*, 1998).

APC is expressed in both epithelial and mesenchymal lineages and particularly high expression has been noted in the central nervous system (CNS) (Smith *et al.*, 1993; Bhat *et al.*, 1994; Miyashiro *et al.*, 1995; Senda *et al.*, 1996; Midgley *et al.*, 1997; Senda *et al.*, 1998).

Bhat and Co-workers (1994) demonstrated high levels of *APC* mRNA throughout the brain of embryonic and early postnatal rats which, with the exception of the olfactory bulb, decreased 6 weeks after birth and remained low in adult rats. Relatively high levels of mRNA were noted in the hippocampus and cerebellum and in layers containing newly formed postmitotic neurones, lower levels were observed in proliferative zones where neurogenesis occurs. The expression pattern seen throughout the developing tissue suggested that APC might contribute to suppressing neuronal proliferation during brain development.

Senda and Co-workers (1998) analysed the localisation of Apc in the mouse CNS. This study showed *Apc* expression in neurones and nerve fibres in some but not all nervous tissues. Apc was shown to be concentrated in the terminal plexus of the basket cell fibres around Purkinje cells. This is also a location of postsynaptic densities (PSD) which are proposed to have several functions, including the stabilisation of synaptic junctions and regulation of neurotransmitter receptors (reviewed in Kennedy *et al.*, 1993). APC was shown to co-localise at these sites with β-catenin and the human homolgue of the Drosophila tumour suppressor, Dlg (lethal (1) disc-large) (Lue *et al.*, 1994; Hunt *et al.*, 1996; Makino *et al.*, 1997; discussed in detail in section 2.3.5).The co-localisation of

these proteins to PSD implicates their involvement in signal transduction between basket cell fibres and Purkinje cells.

Apc was also shown to localise to perivascular astrocytic "endfeet". Astrocytes, one of the glial cells, surround and extend along neurones (both their cell bodies and their processes) and themselves extend processes that are expanded into "endfeet" which, linked by cell junction complexes form a sealed barrier at the external surface of the central nervous system. Astrocyte processes and endfeet act as the sites for cell adhesion and intercellular signalling (Kimelberg *et al.*, 1989; Walz and Wuttke, 1989). The localisation of Apc suggests it may have a role in the regulation of intracellular signal transduction and structural maintenance of astrocyte endfeet. β-catenin or DLG did not co-localise with APC in astrocytic processes and endfeet, this may suggest a different mechanism of signal transduction in astrocytes.

In summary investigations of Apc localisation and expression levels within the CNS may suggest a role for Apc in neuronal cell signalling and cellular proliferation in some but not all cell types. APC is a negative regulator of wnt/β-catenin signalling pathway (discussed fully in section 2.6). One may hypothesise that mutations in the *APC* gene, that result in the perturbation of wnt signalling, may lead to aberrant cellular signalling and neuronal proliferation. It is therefore interesting to note that individuals with germline mutations within *APC* are associated with Turcots syndrome, a syndrome characterised by not only colonic adenomas but also neuroepithelial tumours. Futhermore, murine models with targeted disruption of the proto-oncogene *Int-1/Wnt1* that results in the absence of signalling are characterised by severe brain development abnormalities (Thomas and Capecchi 1990).

In numerous mouse tissues *APC* expression has been shown to vary in its subcellular location in individual cell types of the same tissue. The level of expression also varies and studies have shown increased expression in cells where cell replication has ceased and terminal differentiation is established (Miyashiro *et al.*, 1995; Midgley *et al.*, 1997). This is shown to be the case in enterocytes of both the small and large intestine, with *Apc* expression increasing from the base of crypts to the luminal surface (as cells become

differentiated and no longer undergo mitosis). Diffuse cytoplasmic staining seen in the lower portion of the crypts becomes localised to lateral and sub-apical cell membranes. In the small intestine, enterocytes that have migrated to the luminal surface show strong accentuation of staining along their apical surfaces, this apical staining is maintained as cells migrate up the villi. Apc co-localises with both α- and β-catenin at the lateral membranes but this co-localisation does not occur at apical surfaces (Miyashiro *et al.*, 1995; Senda *et al.*, 1996) supporting additional functions of APC independent of its association with β-catenin. Interestingly, the apical region of epithelial cells contain ends of microtubules. APC has been shown to promote microtubule assembly *in vitro* and co-localisation studies reveal that APC is localised in clusters near ends of microtubules that protrude into actively migrating membrane structures (Munemitsu *et al.*, 1994; Smith *et al.*, 1994; Näthke *et al.*, 1996; discussed in section 2.3.4). Therefore, the localisation of APC to the apical surface of the cell may reflect a role in cell migration. The lateral and sub-apical membranes are typically sites for adhesion junctions; junctions that are required not only for intercellular adhesion, but also for intercellular signalling and cell polarity. The role of APC in cell adhesion is discussed in section 2.7.

Within murine lung tissue, APC expression is described as showing a diffuse cytoplasmic pattern with linear membranous staining of the bronchial epithelial cells (Midgley *et al.*, 1997).

Immunohistochemical studies have generally described *APC* expression as either diffuse or punctate in the cytoplasm with occasional nuclear staining (Miyashiro *et al.*, 1995; Midgley *et al.*, 1997). From early immunohistochemical studies, it was unclear whether nuclear staining was specific. APC nuclear localisation has since been confirmed by cell fractionation techniques and experiments have demonstrated that the ability to clearly see the nuclear component of APC is highly dependent on protocol details e.g. the use of paraformaldehyde versus formaldehyde in cell fixation (Neufeld and White, 1997). Two putative nuclear localisation sequences have been identified within APC. These sequences begin at amino acids 1773 and 2054. The significance of nuclear localisation of APC has not been determined although current knowledge on the interaction of APC with other

cellular proteins does raise some interesting proposals. APC is a negative regulator of the wnt signalling pathway, functioning to negatively regulate the level of β-catenin in the cytoplasm (reviewed in section 2.6). The association of β-catenin with APC, Axin and glycogen synthase kinase 3β leads to β-catenin phosphorylation and the subsequent degradation by the ubiquitin-proteosome system. The accumulation of free β-catenin (not in a complex with cadherin) in the cytoplasm results in translocation to the nucleus where it associates with members of the lymphocyte enhancer binding factor or T-cell factor (LEF-1/TCF) family of transcription factors. As β-catenin and APC interact in the cytoplasm and both proteins have been shown to localise to the nucleus, this raises the possibility of interactions between these two proteins occurring in the nucleus. It is unknown whether APC can interact with β-catenin whilst it is in a complex with LEF-1/ TCF, or if the β-catenin complex is stable, but one could speculate that APC may further negatively regulate wnt signalling at the nuclear level.

β-catenin can be imported into the nucleus and evidence suggests it may also be exported from the nucleus (Fagotto *et al.*, 1998; Prieve and Waterman, 1999), the mechanism of translocation and exportation are unknown as is the nuclear function of uncomplexed β-catenin, it may be possible that APC could have a role in these processes.

The detection of APC in the nucleus may suggest a functional role in mitosis (see section 2.3.4). Studies have shown interactions with microtubules, a key component of the mitotic spindle, and the EB1 family of proteins, proteins that in yeast have been suggested to function as a cell checkpoint mechanism prior to mitosis, whilst immunohistochemical analysis suggests that APC is a negative regulator of cell proliferation.

2.2 Alternative splicing of the *APC* gene

The *APC* gene was originally reported to contain an open reading frame that encoded a 2843 amino acid sequence. The gene was described as containing 15 exons, exon 1 contained a translation initiation codon and exon 9 was shown to be alternatively spliced and was designated exon 9 and 9A (Groden *et al.*, 1991). To date 5 additional exons have been identified. Four exons are located 5' of exon 1 and are designated (from 5'-3') 0.3, BS (for brain specific), 0.1 and 0.2 (Horii *et al.*, 1993; Thliveris *et al.*, 1994). Transcripts containing the BS exon were originally thought to be only expressed in the brain, this is now known not to be the case. One exon has been identified 3' to exon 1 and is designated exon 10A. Exons 0.3, BS, 0.1, 0.2, 9 and 10A are alternatively spliced to produce at least 16 different transcripts (Santoro and Groden 1997). As with exon 1, exons 0.3, BS and 0.2 also contain a translation initiation codon that is inframe with exon 2. Exons are depicted in figure 5a.

Transcripts that contain exons upstream of exon 1 have been identified, some transcripts encode exon 1 whilst others do not. Sequence analysis revealed a stop codon located 6 base pairs 5' to the exon 1 initiation codon. It is predicted that translation initiated 5' to exon 1 will cease at the stop codon in exon 1 and be reinitiated by the exon 1 initiation codon. Thus only transcripts lacking exon 1 have the ability to encode differentially expressed APC peptides with novel amino terminal domains (Thliveris *et al.*, 1994; Santoro and Groden, 1997; Pyles *et al.*, 1998).

In transcripts identified with exons upstream of exon 1 all were shown to exist in APC isoforms both with and without exon 9 in all cell lines and tissues examined (Santoro and Groden, 1997). This study did not examine the presence of the alternatively expressed exon 10A.

Elevated levels of transcripts lacking exon 1 have been identified in postmitotic and terminally differentiated tissues and can be induced in cell culture systems where cells are driven to differentiate (Santoro and Groden, 1997). This may suggest that different APC isoforms may be involved in the physiology of cell cycle cessation.

Analysis of the DNA and amino acid sequence of *APC* revealed that exon 1 encodes the amino terminus of the oligomerisation domain (discussed in detail in section 2.3.1). The predicted structure of proteins encoded from transcripts without exon 1 suggests that these proteins would lose their ability to dimerise or interact with other proteins through this domain. Further to this, Pyles and co-workers (1998) have shown that APC protein encoded by transcripts lacking exon 1 can not dimerise to APC isoforms that contain exon 1 encoded sequence.

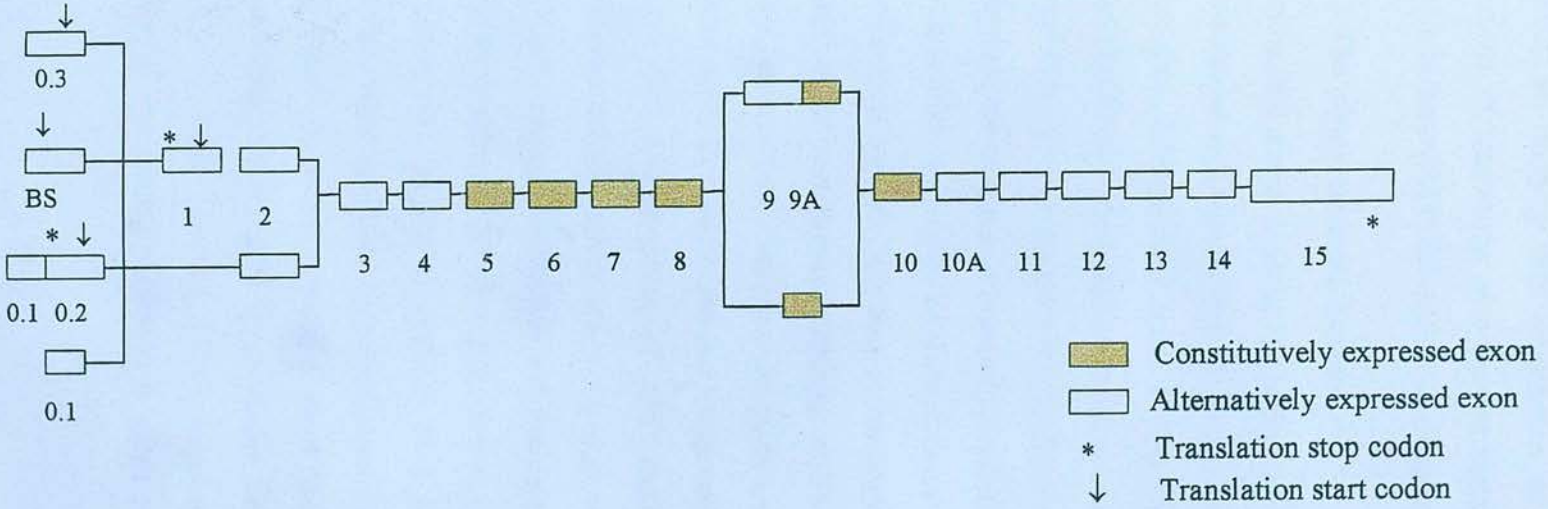
In summary, it is likely that different isoforms of APC specify various cellular functions thereby adding a new dimension to the cellular role of APC.

Figure 5: The *APC* gene: exons and protein structure

(a) Twenty-one exons have currently been identified within the *APC* gene. These exons are alternatively spliced to produce numerous transcripts. Translation initiation codons have been identified in four exons and translation stop codons in 3 exons.

(b) A schematic representation of the APC protein highlighting sites of protein/protein interaction. A number of domains have been identified throughout the APC protein. Some of these domains have been shown to mediate interactions with other cellular proteins whilst other protein interaction domains are still putative (see main text for details).

a :Exons identified within the *APC* gene.



b : A schematic representation of the APC protein highlighting sites of protein/protein interaction.

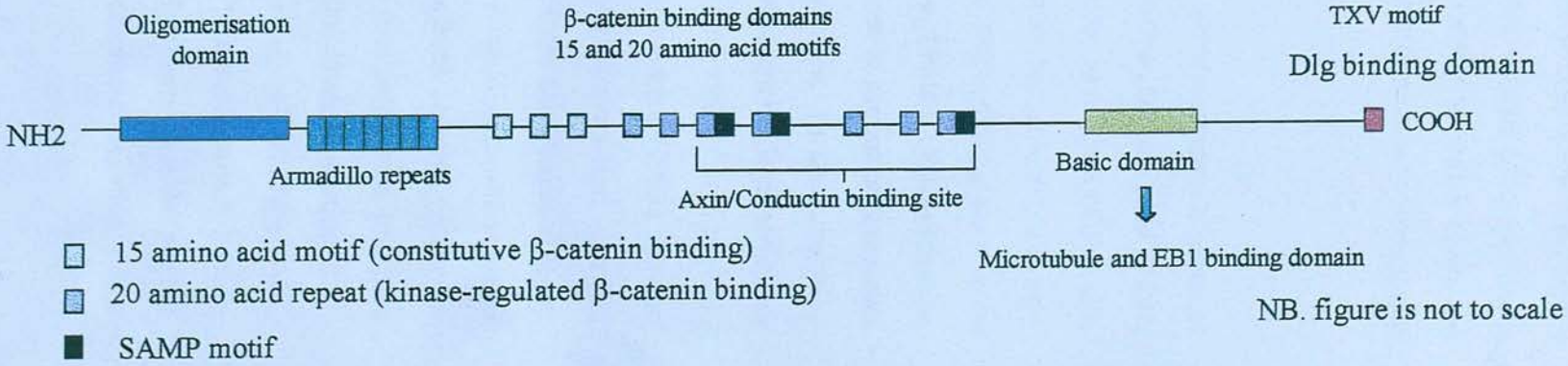


Figure 5: The *APC* gene: exons and protein structure

2.3 Identified functional domains

Nucleotide and amino acid analysis and immunoprecipitation studies have identified a number of domains in APC that are thought to interact with various proteins. Some domains are binding sites of proteins of known function whilst in other cases the functions of interacting proteins are unclear. Each identified domain and interacting protein is discussed in this section and depicted in figure 5b.

2.3.1 The oligomerisation domain

The first 900 amino acids at the amino terminus of the APC protein contain a series of heptad amino acid repeats that are involved in APC oligomerisation (Groden *et al.*, 1991; Joslyn *et al.*, 1991; Kinzler *et al.*, 1991a). These repeat sequences have hydrophobic side chains that form a hydrophobic streak on one side of the protein helix thereby allowing hydrophobic interactions between similar streaks on other helices promoting association of the two helices into a coiled-coil (Traub 1985; Bourne 1991a; 1991b). Oligomerisation may occur between two of the same type of proteins; homo-dimerisation or between two different proteins with oligomerisation domains; hetero-dimerisation. To date the APC protein has not been shown to interact with any other proteins through its oligomerisation domain, however APC has been shown to homo-dimerise (Su *et al.*, 1993).

Using *in vitro* expression and immunoprecipitation Su and co-workers (1993) showed that the first 171 residues of APC are sufficient for homo-dimerisation. Further experiments using lymphoblastoid cell lines, that co-express wildtype and truncated forms of APC, revealed that wildtype and truncated protein products could homo-dimerise *in vitro*. The majority of germline and somatic mutations that have been detected within the *APC* gene lead to the production of a truncated protein product, these products commonly have an intact oligomerisation domain (discussed in detail in section 2.4). Thus, these truncated products have the ability to dimerise to wildtype APC demonstrating the potential for dominant interference by the truncated mutant protein (Joslyn *et al., 1993;* Su *et al., 1993*). Interference of the mutant protein upon the normal function of the wildtype protein is often referred to as a dominant negative effect

(Herskowitz 1987). The hypothesis behind the dominant negative effect is that protein oligomerisation creates a functional complex, if the protein dimerises with a mutant protein then the complex may not be functional. Thus, the wildtype protein is essentially being inactivated.

2.3.2 Armadillo repeats

The armadillo motif is a 42 amino acid sequence that was originally identified in the *Drosophila* segment polarity gene product armadillo, the motif is repeated 12 times in tandem in the central region of the protein (Riggleman *et al.*, 1989). The armadillo motif has also been identified in the vertebrate homologues β-catenin (McCrea *et al.*, 1991) and plakoglobin (γ- catenin) (Franke *et al.*, 1989) and in a number of other proteins with diverse cellular functions. Disparate activities, including nuclear transport, cell adhesion, cell cycle control and microtubule stability, have been ascribed to the armadillo containing proteins and it is thought that the armadillo domain represents a site of protein-protein interaction (reviewed in Hatzfeld, 1999). The specificity of protein-protein interactions could be determined by either individual or small groups of motifs as each motif within a protein is generally only 30% identical. Individual motifs are highly conserved during evolution presumably reflecting well-conserved interactions with target proteins (Peifer and Wieschaus 1990).

APC also encodes the armadillo motif, which is repeated 7 times in tandem; this domain is situated at amino acids 453-767 (Kinzler *et al.*, 1991a; Groden *et al.*, 1991; Peifer *et al.*, 1994). APC interacts directly with the armadillo repeat domain of β-catenin although this interaction is not through APC's armadillo repeat domain (Su *et al.*, 1993; Rubinfeld *et al.*, 1993). APC does not associate directly or indirectly with cadherins (Rubinfeld *et al.*, 1993; 1995) and to date no proteins have been identified that bind to APC's armadillo domain and so the function of APC armadillo motifs remain elusive.

2.3.3 β-catenin binding sites

β-catenin binding sites have been identified in the APC protein. The interaction of APC and β-catenin was reported by two groups who used immunoprecipitation to identify cellular proteins that interacted with APC (Su *et al.*, 1993; Rubinfeld *et al.*, 1993). Two proteins, p95 (95 kDa) and p100 (100 kDa) were co-precipitated with APC and were subsequently identified as β-catenin and α-catenin respectively. APC was found to preferentially interact with β-catenin. These interactions were through amino acids 1014 to 1210 of the APC protein, this region harbours 3 imperfect repeats of 15 amino acids. Each 15 amino acid motif was shown to be capable of binding both α-catenin and β-catenin.

Rubinfeld and co-workers (1993) identified a second β-catenin binding motif, a 5 amino acid sequence, serine-leucine-serine-serine-leucine, that is also present in the β-catenin binding domain of E-cadherin. This sequence formed part of the of the 20 amino acids motif described by Groden *et al.*, (1991), which is repeated 7 times within amino acids 1034 and 2130 of the APC protein.

The 15 and 20 amino acid motifs share some sequence homology but are distinct sequences. Experiments have demonstrated that in order for the 20 amino acid repeat domain to bind β-catenin this region must be phosphorylated (Rubinfeld *et al.*, 1996), discussed fully in section 2.6. Hence this β-catenin binding domain is often referred to as the kinase-regulated binding domain whereas the 15 amino acid repeat domain is referred to as the constitutive binding domain. The kinase regulated binding domain has been shown to down regulate β-catenin levels (Rubinfeld *et al.*, 1993; Su *et al.*, 1993; Munemitsu *et al.*, 1995). It is currently unclear how this region leads to this function but it is interesting to note that interspersed between motifs are SAMP (Ser-Ala-Met-Pro) sequences, these sequences are situated downstream of motif number 3, 4 and 7. The protein Axin has been identified and implicated in the negative regulation of β-catenin, evidence suggests that Axin interacts with APC through APC's latter five 20 amino acid motifs and that this interaction requires SAMP sequences (Behrens *et al.*, 1998; Hart *et*

al., 1998; Nakamura *et al.*, 1998). Interactions between APC, Axin and β-catenin are discussed fully in section 2.6.6

The identification of distinct β-catenin binding domains in APC raised considerable interest and investigators explored possible cellular functions of the APC and β-catenin interaction. Two major cellular functions were determined 1) the interaction of APC and β-catenin in the wnt signalling pathway and 2) the influence of APC on β-catenin mediated cell adhesion. The functional role of APC in these systems is discussed in detail in sections 2.6 and 2.7.

2.3.4 The basic domain and EB1 protein interactions

The carboxyl-terminal third of the APC protein contains a stretch of approximately 200 amino acids that is referred to as the basic domain. This domain is defined as amino acids 2200-2400 (Groden *et al.*, 1991). Analysis of the basic domain reveals a sequence enriched in serine-proline and threonine-proline motifs, motifs that have been shown to exist in phosphorylated and unphosphorylated states (Bhattacharya and Boman, 1995). The basic domain has amino acid sequence homology to many microtubule associated proteins (MAPs), these proteins play a role in modulating the functions of microtubules. This may suggest that APC may also have a function in microtubule dynamics. Recent investigations have supported this theory.

Microtubule associated proteins through the modulation of microtubules (MTs) influence cell polarity, the internal and external shape of the cell and are also involved in cell migration (reviewed in Wadsworth 1999; Waterman-Stoner and Salmon 1999). Microtubules also play a crucial role in chromosome and spindle movement in meiosis and mitosis (Hyman and Karsenti 1996; Vernos and Karsenti 1996) and in the transport of vesicles and movement of organelles (reviewed in Lane and Allan 1998).

Investigations into the localisation of APC revealed that full-length APC protein and fragments containing the carboxyl-terminal of APC associate along the length of cytoplasmic microtubules *in vivo*. APC protein lacking the carboxyl terminal region did

not result in this filamentous pattern of staining and was found to be distributed in deposits throughout the cell cytoplasm (Munemitsu *et al.*, 1994; Smith *et al.*, 1994). Wildtype and carboxyl protein fragments were also shown to promote the assembly of microtubules *in vitro*. These results suggested that mutations within *APC* that result in the truncation of APC prior to the basic domain could have an effect on important microtubule processes. As these experiments were carried out on cell lines that were transfected with plasmids overexpressing fragments of *APC*, the pattern of protein expression might not reflect the localisation of endogenous levels of APC. Näthke and co-workers (1996) examined the localisation of endogenous APC in MDCK and IEC-6 cells and determined that APC protein is localised in punctate clusters near the ends of microtubules that protrude into actively migrating membrane structures. These finding support previous experiments showing that APC protein associates with microtubules and further suggests a role in microtubule dependent cell migration. It could be hypothesised that APC is localised to sites of cell migration and is involved in the bundling of microtubules, forming structures that stabilise the direction of migration. Alternatively, the increased level of APC detected at these sites could lead to a reduction in cell-cell adhesion due to increased turnover of β-catenin leading to a reduction in cadherin/catenin complexes (discussed in section 2.7). Further to this, Näthke and co-workers detected APC clusters at the outer boundary of cell-cell contacts where membranes are moving past each other. The importance of APC in cell migration has recently been elucidated by the analysis of murine models. Mice that harbour a germline mutation at codon 850, a mutation that results in the production of a truncated protein product lacking the basic domain, have been shown to exert a dominant negative effect on the migration of enterocytes in colonic crypts (Mahmoud *et al.*, 1997).

2.3.4.1 APC, EB1 and microtubules

Using the yeast two-hybrid system with a fragment of *APC* encompassing amino acids 2167-2843 as bait, Su *et al.*, (1995) identified a cDNA encoding a novel protein termed EB1 that could associate with APC *in vivo* and *in vitro*. The functions of EB1 and EB1-

APC interactions are currently unclear.

EB1 has been found to localise, in mammalian cells, to sections of microtubule filaments, in particular to leading edges of cell membranes and the periphery of stained fibres. Perinuclear staining was present with intensive staining at the microtubule organising centres (MTOC). During mitosis, EB1 localised to spindle microtubules and centrosomes and was strongly associated with midbody microtubules during cytokinesis. Analysis of the localisation of EB1 in the colorectal carcinoma cell line SW480, which encodes a truncated APC protein lacking the EB1 binding domain, revealed that EB1 localisation to microtubules and the centrosomes was preserved. This demonstrated that cellular distribution of EB1 does not depend on the carboxyl terminus of APC (Morrison *et al.*, 1998; Berrueta *et al.*, 1998).

Shortly after the identification of *EB1*, Renner *et al* (1997) published details of a gene *RP1* that was shown to be rapidly up regulated following T-cell activation and that was expressed in cell lines, derived from a spectrum of tumours, that were rapidly proliferating but not cells blocked in G_1. The expression pattern of *RP1* in activated T cells, i.e. rapid up-regulation peaking after 4 hours prior to a decline in expression, suggested that *RP1* might be a member of the immediate-early gene family. Immediate-early genes and the proteins encoded by these genes can act as pleiotrophic regulators of cell activation, and include transcription factors and proteins involved in signal transduction cascades (Kelly and Siebenkist 1995). *RP1* had significant homology to *EB1* and experiments revealed specific binding to wildtype but not to truncated APC protein. Two further genes with high homology to *RP1* were identified using the technique of RACE (rapid amplification of cDNA ends), these genes were designated *RP2* and *RP3*. Together these genes presently constitute the *EB1* gene family.

Homologues of *EB1/RP* have been identified in mice and rats at both the nucleotide and protein level, and in yeast homologues *Mal3* (*Schizosaccharomyces pombe*) and *Bim1*p (*Saccharomyces cerevisiae*) have recently been identified (Schwartz *et al.*, 1997; Beinhauer *et al.*, 1997). The human EB/RP proteins and the yeast Mal3/Bim1p proteins have been shown to bind to cytoplasmic and nuclear microtubules *in vivo* suggesting

conserved function (Schwartz *et al.*, 1997; Beinhauer *et al.*, 1997; Berrueta *et al.*, 1998; Morrison *et al.*, 1998; Juwana *et al.*, 1999). The analysis of yeast cells harbouring deleted *Bim1p* or *Mal3* reveal that these proteins play a critical role in microtubule dynamics and may function as a cell checkpoint mechanism during mitosis (Schwartz *et al.*, 1997; Beinhauer *et al.*, 1997; Muhua *et al.*, 1998). *Mal3* deletion mutants showed condensed chromosomes suggesting a delay in mitosis whilst in *Bim1p* deleted cells the mitotic spindle was abnormally short. Furthermore, in cells harbouring a mutation in *act*5 (gene encoding for the motor protein dynein) that results in the mis-positioning of the nucleus, *Bim1*p was shown to block cytokinesis. No yeast homologue of APC has been identified therefore, these experiments can not be expanded to examine the role of APC in yeast cytokinesis.

In summary the EB/RP family have been shown to interact with the C-terminal domain of the APC protein however at present the significance of this interaction is unclear.

2.3.5 Disc large proteins

The tumour suppressor gene *dlg*-A (lethal (1) discs-large) was originally identified in studies of *Drosophila*. Germline mutations within *dlg-A* resulted in neoplastic overgrowth of the imaginal disc epithelium, loss of apical-basolateral polarity and disruption of normal cell-cell adhesion (Stewart *et al.*, 1972; Woods and Bryant, 1989). This phenotype suggested that the function of wildtype dlg-A protein may be concerned with maintaining normal epithelial structure and possibly with regulating cell division and differentiation (Woods and Bryant, 1991).

Dlg-A is the founding member of an expanding group of proteins referred to as the MAGUK (membrane-associated guanylate kinase) family. MAGUK proteins are typically localised to various cell junctions and play an important role in maintaining junction structure and signalling properties. These functions are achieved through three distinct structural domains that mediate specific protein/protein interactions. The structural domains of the protein are: an N-terminal segment comprising one or more discs-large

homologous regions (DHRs); a Src oncogene homology motif 3 (SH3); and a C-terminal domain with homology to guanylate kinase (GUK).

Homologues of *Drosophila* dlg–A have been identified; h-DLG and NE-DLG (highly expressed in neuronal and endocrine tissues) are human homologues (Lue *et al.*, 1994; Makino *et al.*, 1997); m-Dlg and synapse-associated protein 97 kDa (Sap-97) have been identified as mouse and rat homologues respectively (Muller *et al.*, 1995; Lin *et al.*, 1997).

The analysis of vertebrate homologues show that Dlg is a ubiquitously expressed, with highest levels noted in the central nervous system (CNS) (Muller *et al.*, 1995; Lin *et al.*, 1997; Makino *et al.*, 1997). The protein localises to regions of cell-cell contact and in stratified epithelium has been shown to be specific to non-proliferating cells. Within the CNS, is concentrated at synaptic junctions (within postsynaptic densities) in dendrites and cell bodies. Postsynaptic densities (PSDs) function to stabilise synaptic junctions and regulate neurotransmitter receptors (reviewed in Kennedy *et al.*, 1993). MAGUKs localised to PSDs, including SAP-97, have been shown to bind N-methyl-D-aspartate (NMDA) receptors, Shaker-type potassium channels, and brain nitric oxide synthase through the interaction of the DHR region in the MAGUK protein and the (T/S)XV (threonine or serine, any amino acid, valine) motif located at the C-terminus of the target proteins (Niethammer *et al.*, 1996; Doyle *et al.*, 1996; Brenman and Bredt, 1997). The human homologues of dlg interact with an (T/S)XV motif located at the C-terminal 15 amino acids of the APC protein via the DHR region (Matsumine *et al.*, 1996; Makino *et al.*, 1997).

The function of APC/dlg protein interactions is currently unknown. Both proteins function to suppress tumourigenesis. They co-localise to regions associated with adherens junctions (Matsumine *et al.*, 1996; Midgley *et al.*, 1997) and are both likely to interact with a number of other proteins associated with cell signalling and junction formation. One could hypothesise that dlg and other MAGUK proteins may compete for the binding of APC's carboxyl terminal (T/S)XV motif, and its localisation to cell-cell contacts may suggest interaction with proteins such as the zonula occludentes proteins (zo-1, zo-2 and

zo-3). These proteins are associated with intercellular adhesion being localised at tight junctions, adherens junctions and desmosomes where they appear to function as cross-linkers between occludin and actin filaments in epithelial/endothelial cells or between α-catenin and actin filaments in non-epithelial/endothelial cells (Itoh *et al.*, 1997; Haskins *et al.*, 1998; Itoh *et al.*, 1999).

Within the CNS, one may speculate that APC could compete with NMDA receptors and shaker-type potassium channels for binding of dlg and regulate the clustering of receptors and channels thereby influencing their functions.

2.4 Mutations in the *APC* gene

Over 1,000 mutations have been identified in the *APC* gene, the majority from the analysis of FAP and sporadic colorectal cancer cases. Somatic mutations have been described in other neoplasms; pancreas (Yashima *et al.*, 1994; Horri *et al.*, 1992a), stomach (Horri *et al.*, 1992b; Nakatsuru *et al.*, 1992; 1993; Tamura *et al.*, 1994), oesophagus (Powell *et al.*, 1994; Gonzalez *et al.*, 1997) liver (Oda *et al.*, 1996; Imai *et al.*, 1997) and prostate (Watanabe *et al.*, 1996). The frequency of somatic *APC* mutations within oesophageal, prostate and pancreatic cancer is less than 10%, stomach cancers is approximately 20% and hepatoblastomas 61%.

Of the mutations identified, the vast majority (>98%) lead to the premature truncation of the APC protein (See Laurent-Puig *et al.*, 1998). More than 98% of all mutations are frameshift mutations or nonsense mutations (two thirds are frameshift mutations whilst one third are nonsense mutations).

In colorectal cancer, missense point mutations were initially thought to be extremely rare however recent reports have described germline missense mutations that confer an increased risk to colorectal cancer (White *et al.*, 1996; Laken *et al.*, 1997; Frayling *et al.*, 1998). The mutations detected in the *APC* gene in hepatoblastomas differs from other neoplasms with the majority of mutations being missense mutations (90%) with only 10% of cases featuring a frameshift mutation (Oda *et al.*, 1996; See Laurent-Puig *et al.*, 1998).

Frameshift mutations, insertions or deletions of one or more base pairs, may result from chemical mutagens that intercalate in DNA. The effect of these mutations is to alter the reading frame in translation unless an integral multiple of 3 is inserted or deleted. Missense and nonsense mutations occur due to the substitution of one base for another. These mutations may result from spontaneous deamination of cytosine and failure of DNA repair systems to correct the error leading to C/G to T/A nucleotide transition, both mutations could occur at an equal frequency. However, 98% of mutations are C-T transitions and only 2% are G-A transitions. This bias is thought to be due to the fact that most transitions occur at CGA codons and a C-T transition results in a nonsense mutation TGA which signals translation to stop. Thus, this mutation reflects selection pressure for making a truncated protein.

Alternatively, these mutations may result from mutagens that chemically modify the DNA base resulting in an unidentifiable base, which can not be duplicated during replication. DNA polymerase will incorporate a base into the daughter strand, in most instances this base tends to be a thymine. This substitution can result in a change in the amino acid (missense mutation) or a codon that does not encode an amino acid leading to the premature termination of translation (nonsense mutation).

Almost all germline and somatic *APC* mutations occur in the 5' half of the open reading frame (see figure 6). Germline mutations at codon 1061 and 1309 are extremely frequent and account for about 30% of cases. The majority of mutations at these sites involve 5 bp deletions that result in the production of a premature stop codon. Somatic mutations most commonly occur at codons 1309 (7%) and 1450 (8%) and mutations cluster in a region referred as the mutation cluster region (MCR), located between codons 1286 and 1513 (Miyoshi *et al.*, 1992b). The location of these mutation hotspots and the MCR is within the kinase regulated β-catenin binding domain, and these mutations are likely to disrupt the function of this domain. Also, as the majority of mutations result in premature stop codons all downstream domains are usually lost (see figure 6).

Following the analysis of colorectal adenomas arising in FAP patients it has been shown that patients with a germline mutation in the *APC* gene which resides within a small

region (codons 1194-1392), mainly show allelic loss in their colorectal adenomas in contrast to other FAP patients, whose second hit tends to occur by truncating mutations in the mutation cluster region (Lamilum *et al.*, 1999). This study suggests that different *APC* mutations provide cells with different selective advantages.

The location of the *APC* germline mutation is also associated with severity of polyposis and the presence of extracolonic manifestations, namely congenital hypertrophy of the retinal pigment epithelium (CHRPE), desmoid tumours, oesteomas and epidermoid cysts (Nagase *et al.*, 1993; Olschwang *et al.*, 1993; Wallis *et al.*, 1994; Bunyan *et al.*, 1995; Caspari *et al.*, 1995; Dobbie *et al.*, 1996). Different phenotypes may be due to loss of various functional domains, dominant negative effects of wildtype and mutant protein interactions. However, germline deletions of *APC* lead to classical FAP and attenuated FAP, and the extracolonic manifestations, desmoid, oesteomas and epidermoid cysts have all been detected. This suggests that factors additional to the location of the germline mutation are of importance and to this effect. A modifier locus, MOM-1 (modifier of Min) has been mapped in the mouse model of FAP to the region of murine chromosome 4 (Dietrich *et al.*, 1993; MacPhee *et al.*, 1995; Luongo and Dove, 1996; Shoemaker *et al.*, 1996). This murine chromosomal region has synteny to human chromosome 1p35-36 and a study of a large FAP kindred, in which patients harbour the same germ-line mutation but show markedly different disease characteristics, is supportive of a phenotype-modifying locus within this chromosomal region (Dobbie *et al.*, 1997). The secretary type II phospholipase gene (*Pla2g2a*) has been suggested as a candidate gene which may modify the FAP phenotype. This gene is located in mice and humans within the chromosome regions identified as harbouring the modifier locus, MOM-1 (Praml *et al.*, 1995; MacPhee *et al.*, 1995; Tomlinson *et al.*, 1996a). Pla2g2a functions in the cleavage of fatty acids from lipids and is expressed in the small intestinal crypts (McPhee *et al.*, 1995), therefore expression of Pla2g2a is thought to alter the cellular microenvironment within the intestinal crypt. Overexpression of the gene has been shown to reduce tumour multiplicity and size within a murine model of FAP (Cormier *et al.*, 1997). MOM-1 sensitive strains are thought to be null for Pla2g2a protein, C57BL/6 is a

sensitive strain for MOM-1.

The human PLA2GA gene has been sequenced and polymorphic variants detected in a study of FAP patients. No associations were found between severity of polyposis and PLA2GA variants (Tomlinson *et al.*, 1996b) and to date there is no data to support PLA2GA as a modifier of FAP in *Homo sapiens*.

Figure 6: Functional domains within the APC protein and distribution of mutations identified

Analysis of familial and sporadic colorectal cancers and other sporadic tumours has revealed the mutation spectrum that is depicted in figure 6. The majority of mutations identified occur within the 5' portion of the protein and usually result in the production of a truncated protein product. The region most commonly mutated is referred to as the mutation cluster region (MCR). This region spans the kinase-regulated β-catenin binding domain. Tumours that produce APC protein that encompasses the kinase regulated β-catenin binding domain are extremely rare. From this evidence, it appears that this domain is crucial in tumour suppression. The kinase regulated β-catenin domain has been shown to down regulate β-catenin levels. Accumulation of β-catenin within the cytoplasm results in the transcriptional activation of target genes of the wnt signalling pathway. A number of these genes have already been described as oncogenes. See section 2.6 for full discussion.

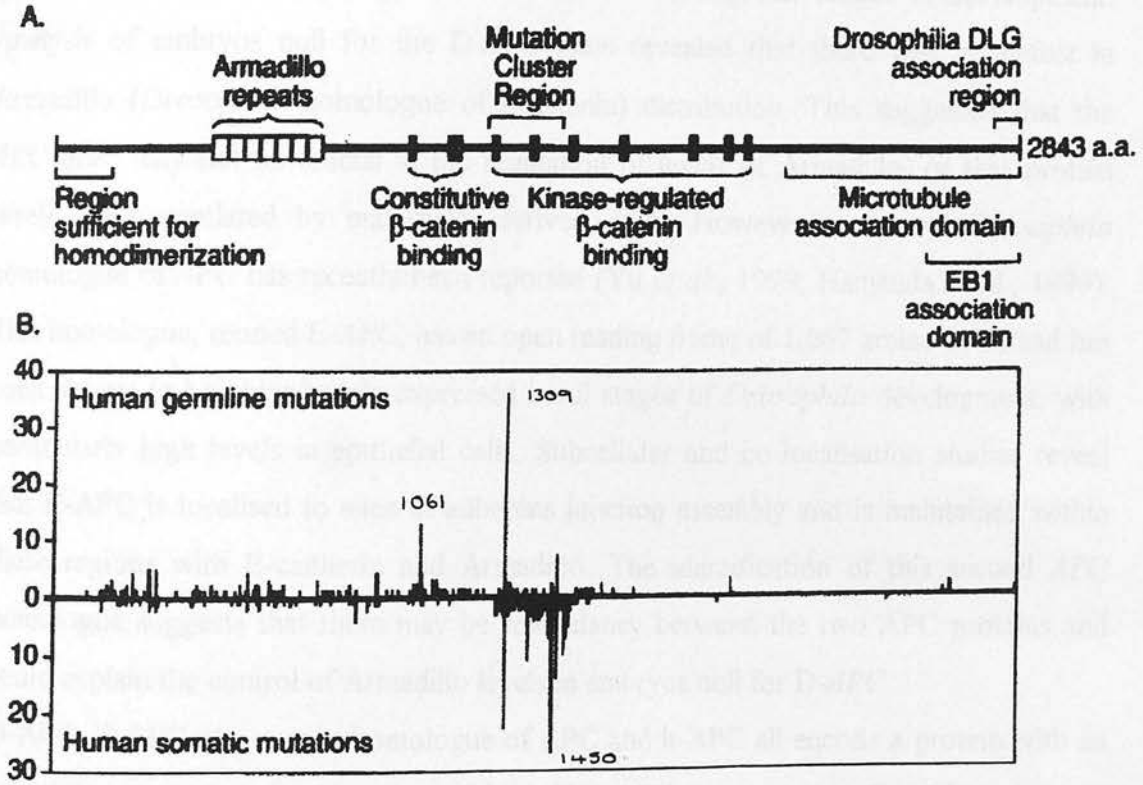
A: A linear diagram showing identified functional domains in the APC protein.

B: Distribution of germline and somatic mutations identified in the *APC* gene relative to protein domains shown in 6a.

Figure: 6

A: A linear diagram showing identified functional domains within the APC protein.

B: Distribution of germline and somatic mutations identified in the *APC* gene relative to protein domains shown in 6a.



2.5 Identification and characterisation of *APC* homologues

The *APC* gene has been shown to exhibit a high degree of homology across species (>86%). Homologues of the human *APC* gene (h-*APC*) have been identified in mouse (Su *et al.*, 1992), *Drosophila* (Hayashi *et al.*, 1997; Yu *et al.*, 1999; Hamanda *et al.*, 1999), *Xenopus* (Vleminckx *et al.*, 1997) and *C.elegans* (Rocheleau *et al.*, 1997). Recently, new homologues of h-*APC* gene have been reported in man (van Ets *et al.*, 1998; Nakagawa *et al.*, 1998).

Hayashi and coworkers (1997) identified the first *Drosophila* homologue (D-APC). Investigations revealed a 2416 amino acid protein that was expressed at low levels in the developing embryo and at high levels in the CNS during late stages of development. Analysis of embryos null for the D-*APC* gene revealed that there was no defect in Armadillo (*Drosophila* homologue of β-catenin) distribution. This suggested that the *APC* gene may not be crucial in the regulation of levels of Armadillo, or that protein levels were regulated by maternally derived APC. However, a second *Drosophila* homologue of *APC* has recently been reported (Yu *et al.*, 1999; Hamanda *et al.*, 1999). This homologue, termed E-*APC*, has an open reading frame of 1,067 amino acids and has been shown to be ubiquitously expressed in all stages of *Drosophila* development, with particularly high levels in epithelial cells. Subcellular and co-localisation studies reveal that E-APC is localised to sites of adherens junction assembly and is maintained within these regions with E-cadherin and Armadillo. The identification of this second *APC* homologue suggests that there may be redundancy between the two APC proteins and could explain the control of Armadillo levels in embryos null for D-*APC*.

D-APC, E-APC, the murine homologue of APC and h-APC all encode a protein with an armadillo repeat domain, 15 amino acid motif (h-APC seven repeats, D-APC one repeat and E-APC two repeats), the 20 amino acid motifs (D-APC and E-APC encode five repeats whereas h-APC encodes seven repeats). Functional analysis revealed that D-APC and E-APC associate with β-catenin *in vitro*, further experiments have shown that E-APC also interacts with the *Drosophila* homologue of GSK-3β (sgg) *in vitro*. Transfection of D-*APC* into an *APC* null colon carcinoma cell line showed that D-APC

could down regulate the intracellular concentration of β-catenin. Analysis of these binding domains suggests that these two *Drosophila* homologues have similar functions to those carried out by h-APC and murine APC.

D-APC, like its human and murine counter parts, also contains a basic domain which is thought to be capable of binding microtubules. However, the carboxyl terminus of the D-APC lacks the X (S/T)XV motif, a motif previously been shown to be required for the interaction of APC with dlg. E-APC also diversifies at this point from human and mouse homologues and initial analysis suggests that in *Drosophila* APC will not bind microtubules or dlg-A.

The *APC* homologue that has been identified in *Xenopus*, X*APC* (Vleminckx *et al.*, 1997) has been shown to be found primarily in a complex with β-catenin. Analysis of the overexpression of X*APC* in ventral blastomeres did not result in changes in β-catenin levels, but resulted in a phenotype, development of a second axis, which is typical of overexpression of the wnt signalling pathway. This study suggested that XAPC might have a positive signalling role in the wnt signalling pathway. Analysis of the *C.elegans* APC homologue, APR-1, support studies in *Xenopus* (Rocheleau *et al.*, 1997), whereas experiments in other species had indicated that APC negatively regulated the wnt signalling pathway, discussed in detail in section 2.6.

Two recent reports have been published which describe the identification and initial characterisation of two new APC protein homologues in *homosapiens*. *APC2* was described by van Ets and co-workers (1998) whilst *APCL* was described by Nakagawa and colleges (1998). Both *APC2* and *APCL* map to human chromosome 19p13.3.

APC2 is a 2274 amino acid protein, whilst APCL encodes a protein of 2303 amino acids. It is likely that these two proteins are isoforms. The two homologues show different expression patterns with APC2 being ubiquitously expressed, with highest levels detected within the CNS, whereas APCL2 is described as brain specific.

Analysis of the two homologues revealed that both encode the same protein binding domains. As detected in the original human APC protein (h-APC) the protein structure of the homologues consists of an N-terminal oligomerisation domain and an armadillo

domain (which is 76% homologous to h-APC). Downstream of these domains in h-APC lies the constitutive β-catenin binding domain encoded by 3 repeats of a 15 amino acid motif. This domain was not present within the human APC homologues. The kinase-regulated β-catenin binding domain, consisting of 7 repeats of a 20 amino acid motif identified in h-APC, differs in the homologues in that only 5 repeats are present. Secondly, this domain varies as only 2 SAMP repeats are interspersed thoughout these repeats as opposed to 3 SAMP repeats. Analysis of the carboxyl terminus of the homologues revealed that the X (S/T)XV motif, a motif previously shown to be required for the interaction of h-APC with DLG, is not present. It is interesting to note that the structure of the APC homologues is extremely similar to the *Drosophila* homologues of APC, containing only 5 repeats of the 20-amino acid motifs and also lacking the DLG binding region (Hayashi *et al.*, 1997; Yu *et al.*, 1999).

Functional analysis of homologues have shown interactions with β-catenin through the kinase-regulated β-catenin binding domain (20 amino acid repeats) and that through this interaction APC homologues decrease the intracellular concentration of β-catenin.

APC-2 interacts with conductin (homologue of Axin) and both APC homologues inhibit β-catenin/TCF mediated wnt signalling *in vitro*.

The identification of APC homologues within species adds a new dimension to all previous investigations. It is likely functional redundancy exists between proteins and further research is now need to determine the unique functions of these genes within the cell.

2.6 APC and the wnt signalling pathway

The wnt signalling pathway is required for a diverse range of biological processes such as embryonic patterning, determination of cell fate and cell proliferation (reviewed in Wodarz and Nusse, 1998; Dierick and Bejsovec 1999). This diversity is reflected in the complex nature of the pathway. The current model of wnt signalling via the β-catenin mediated pathway suggests that in the absence of a wnt signal, the serine/threonine kinase, glycogen synthase kinase-3β (GSK-3β), phosphorylates β-catenin if associated with the multiprotein complex, which contains APC/β-catenin/GSK-3β/Axin and possibly other components that regulate β-catenin phosphorylation. Phosphorylation targets β-catenin for ubiquitin-mediated degradation. When the wnt signalling pathway is activated, via ligand binding to Frizzled receptors, GSK-3β function is inhibited and β-catenin accumulates and translocates to the nucleus where it interacts with LEF-1/TCF transcription factors. Interaction of β-catenin with these transcription factors is thought to relieve LEF-1/TCF transcriptional repression and in turn results in the transactivation of genes that encode LEF-1/TCF binding sites within their promoters. A model for wnt signalling is depicted in figure 7.

Figure 7: Wnt/β-catenin signalling

A simplified canonical pathway of wnt signalling is represented by figure 7.

a) in the presence of wnt ligand

Wnt ligand binds to the Frizzled receptor that transduces a signal to Dishevelled (Dsh). Dsh inhibits the kinase activity of GSK-3β. Inactivation of GSK-3β kinase activity inhibits phosphorylation of β-catenin and other components that are required to negatively regulate β-catenin levels. Thus β-catenin accumulates in the cytoplasm and subsequently translocates to the nucleus. In the nucleus, β-catenin interacts with the transcription factors of the LEF-1/TCF family following the displacement of co-repressors. β-catenin, whilst in a complex with LEF-1/TCF transactivates target genes of the wnt signalling pathway.

b) In the absence of wnt ligand

In the absence of wnt ligand, GSK-3β is active. APC, Axin and β-catenin form a trimeric complex, whilst GSK-3β binds to Axin and phosphorylates all 3 components. Phosphorylated β-catenin is then targeted for degradation by the ubiquitin-proteosome pathway. LEF-1/TCF located in the nucleus is thought to repress expression of target genes in conjunction with co-repressors.

Figure 7: Wnt/β-catenin signalling pathway

Many components of the wnt signalling pathway appear structurally and functionally conserved thus studies from *Drosophila*, *Xenopus*, *Caenorhabditis elegans* and mammals have all contributed to the understanding of this pathway. Identified components of β-catenin mediated wnt signalling in vertebrate and invertebrates are listed in table 3.

Table 3. Wnt signalling components in vertebrates and invertebrates. Multiple homologues have been identified in all cases unless indicated by an asterisk

Generic	Mammalian	*Drosophila*	*C.elegans*	*Xenopus*
wnt	Wnt	D-wnt (D-wnt-1=wg)	Lin-44 MOM-2	Xwnt
Frizzled	Fz	Fz	Lin-17 MOM-5	Non described
Dishevelled (Dsh)	Dvl	Dsh-1*	Putative genes identified	Xdsh
GSK-3	GSK-3 α and β-isoforms	Zw-3/sgg*	Non described	XGSK-3 β-isoforms
Axin	Axin High homology with Axil/Conductin	Daxin	Non described	XAxin
APC	APC	D-APC	APR-1	XAPC
β-catenin	β-catenin	Arm*	WRM-1	β-catenin
TCF	TCF/LEF	dTCF/pangolin	POP-1	XTCF

2.6.1 Wnt ligands

The wnt proteins comprise a highly conserved, multi-member ligand family (Nusse and Varmus, 1982; Nusse *et al.*, 1984; Wainwright *et al.*, 1988; Gavin *et al.*, 1990; Gavin and McMahon, 1992). W*nt* genes encode secreted glycoproteins, usually 350-400 amino acids in length, which contain a conserved pattern of 23-24 cysteine residues in addition to other invariant amino acids. Secreted wnt proteins associate with the extracellular membrane where they may interact with proteoglycans (see below).

2.6.2 Frizzled receptors

Wnt proteins have been shown to bind to the extracellular domain of Frizzled receptors. Genetic screening in *Drosophila* identified the first *Frizzled* (*fz*) gene. Mutations in this gene resulted in a defect in tissue or planar polarity, phenotypically expressed as disruption in the alignment of wing hairs, direction of bristles on the notum and legs and disorientation of the ommatidia comprising the insect compound eye (Zheng *et al.*, 1995). Frizzled proteins are characterised by a large cysteine-rich extracellular domain (CRD) that spans approximately 120 amino acids and contains 10 cysteine residues. The CRD is located at the N-terminal of the protein and this region is the site of wnt binding (Ingham, 1996; Bhanot *et al.*, 1996; Orsulic and Peifer, 1996; Wang *et al.*, 1997b). The CRD is followed by seven putative membrane-spanning domains and a cytoplasmic C-terminal tail. Many Frizzled proteins contain within the carboxyl terminal tail an X(T/SX)V motif, this motif can interact with DHRs (discussed further below and in section 2.3.5) (Chan *et al.*, 1992; Wang *et al.*, 1996).

A number of wnt ligands and Frizzled receptors have been identified and analysis in *Drosophila* has revealed that some receptors are capable of binding different D-wnt ligands and that many members of the D-wnt ligand family will bind to multiple receptors. Endogenous D-wnt expression patterns in many tissues overlap, this raises the possibility that different D-wnt proteins may functionally substitute for each other *in vivo* (Gavin *et al.*, 1990; McMahon *et al*, 1992). Redundancy of receptor binding may be the explanation as to why developmental defects in *Wnt* null mice occur in only a subset of

the regions in which the *Wnt* is expressed (McMahon *et al.*, 1992; Takada *et al.*, 1994; Stark *et al.*, 1994; Parr and McMahon, 1995).

2.6.3 Frizzled related proteins

In addition to the integral membrane Frizzled proteins, *Xenopus* and other vertebrates produce several secreted proteins, called FRPs (Frizzled related proteins) that contain a CRD that is very similar to those in Frizzled proteins. These molecules lack the 7 transmembrane sequence, but encode a stretch of charged residues with homology to netrins; exon guidance proteins found in the developing vertebrate spinal cord (Shirozu *et al.*, 1996; Leyns *et al.*, 1997; Rattner *et al.*, 1997; Wang *et al.*, 1997). FRPs have been shown to complex with wnt ligands and frizzled receptors and *in vitro* experiments reveal that FRPs antagonise wnt signalling (Leyns *et al.*, 1997; Wang *et al.*, 1997; Bafico *et al.*, 1999).

2.6.4 Association with proteoglycans

As previously mentioned, wnt proteins associate with proteoglycans at the cell surface (Schryver *et al.*, 1996). Proteoglycans function as low affinity cell surface receptors for a variety of ligands (Schlessinger *et al.*, 1995). It is thought that proteoglycans may increase the local concentration of wnt proteins leading to increased avidity and/or receptor clustering (Reichsman *et al.*, 1996), or may be required for the transduction of wnt signals.

2.6.5 Transduction of signals via frizzled; G-proteins and Dishevelled

It is likely that the recruitment or release of intracellular signalling molecules transmits wnt signals as no enzymatic motifs have been identified in the Frizzled receptors. Candidate molecules that could interact with Frizzled include heterotrimeric G-proteins and the protein Dishevelled (Dsh).

Heterotrimeric G-proteins have previously been shown to couple numerous receptors that

contain 7 transmembrane sequences to a range of intracellular signalling pathways (reviewed in Knall and Johnson, 1998; Simonds, 1999). Despite only minimal homology between Frizzled proteins and other known G-protein coupled receptors (Mukoyama *et al.*, 1993), there is some evidence to suggest that Frizzled receptors may transmit signals via G-proteins (Slusarski *et al.*, 1997).

The Dsh protein has been shown to be required for wnt signalling in many tissues (Lecuit and Cohen, 1997; Neumann and Cohen, 1997). Several mammalian *Dsh* homologues (*Dvl*) have been identified revealed and an overlapping pattern of expression of proteins have been detected suggesting redundancy in function between these protein homologues (Sussman *et al.*, 1994; Klingensmith *et al.*, 1996; Tsang *et al.*, 1996). Further to this, murine gene knockout models, null for one *Dvl* gene, do not have phenotypes typical of loss of wnt signalling (Sussman *et al.*, 1994; Pizzuti *et al.*, 1996; Lijam *et al.*, 1997).

Amino acid sequence comparison of all known Dsh molecules revealed 3 highly conserved regions thoughout the species: an N-terminal Dsh homology domain, a DHR domain and a DEP (Dsh/egl-10/pleckstin) domain. The Dsh homology domain shares homologies to a newly identified protein, Axin, and these two proteins have recently been shown to interact. Axin is thought to play an important role in the negative regulation of wnt signalling and is discussed further in section 2.6.6 (Zeng *et al.*, 1997; Ikeda *et al.*, 1998; Sakanaka *et al.*, 1998; Behrens *et al.*, 1998). The Dsh domain is now referred to as the DIX domain (Dishevelled and Axin). Schematic representation of Dsh and Axin is depicted in figure 8.

The DHR domain located in the central region of the Dsh/Dvl proteins is also found in a number of other proteins including dlg and zo-1 (Ponting *et al.*, 1997). Structural analysis has shown that the DHR domains within some proteins are able to bind four residue C-terminal peptides with the motif X(S/T)XV (section 2.3.5). The Frizzled receptors have been shown to encode this motif in their cytoplasmic C-terminal tails, however, the Dsh DHR falls into the non-binding class (Songyang *et al.*, 1997; Doyle *et al.*, 1996; Morais Cabral *et al.*, 1996). In keeping with this, experiments have failed to detect interactions between any Frizzled proteins and Dsh proteins (Nusse *et al.*, 1997). At the C-terminal

side of the DHR domain is the DEP domain. This domain is approximately 80 amino acids in length and has been detected in a number of other proteins including the guanine nucleotide dissociation and GTPase-activating protein families (Ponting and Bork, 1996). The molecular mechanism by which Dsh transduces wnt signals is unknown. Dsh is a phosphoprotein that localises predominantly in the cytoplasm of cells (Yanagawa *et al.*, 1995). Exogenous expression of *wnt* ligands in mammalian and *Drosophila in vitro* cell culture systems results in a fraction of Dsh protein localising to the cell membrane (Yanagawa *et al.*, 1995; Steitz *et al.*, 1996). This change in compartment location suggests that the membrane fraction of protein is active in signalling. In *Drosophila,* the Dsh-1 protein that is translocated to the cell membrane is hyperphosphorylated (Yanagawa *et al.*, 1995). Phosphorylation is thought to be by casein kinase II (CKII) as Dsh has been found in a complex with CKII and is phosphorylated by CKII *in vitro* (Willert *et al.*, 1997). It is currently unclear whether phosphorylation of Dsh is a consequence of or is required for wnt signalling.

Recent studies have identified interactions between Axin and Dsh suggesting that Dsh may transduce wnt signals by interacting with the multiprotein complex that is required for β-catenin phosphorylation (discussed further in section 2.6.7).

2.6.6 GSK-3β and wnt signalling

Studies in *Drosophila* revealed that mutations within the gene *zeste white-3* (*zw3*) resulted in phenotypes consistent with the constitutive activation of wnt signalling (Peifer *et al.*, 1994). The β isoform of glycogen synthase kinase-3 (GSK-3β) was identified as the mammalian homologue of *zw3* (Plyte *et al.*, 1992). Glycogen synthase kinase-3 is a widely conserved serine threonine protein kinase, which is constitutively active in unstimulated cells and is inhibited by signal transduction pathways responding to various growth factors in different systems (Kim *et al.*, 1998; Marcus *et al.*, 1998). Data from studies in *Drosophila* and *Xenopus* demonstrate that GSK-3β kinase activity is negatively regulated by and negatively regulates wnt signalling (He *et al.*, 1995; Dominguez *et al.*,

1995; Pierce and Kimelman, 1995).

Wnt signalling is thought to mediate inhibition of GSK-3β kinase activity through *Dishevelled* (*Dvl-1*), although a mechanism has not yet been determined. GSK-3β kinase activity is revoked by incubation with serine/threonine phosphatases and recent investigations demonstrate that Axin interacts with protein phosphatase 2A *in vitro,* suggesting that this may be a mechanism of negative regulation (Hsu *et al.*, 1999), discussed further in section 2.6.7.5.

Analysis of *Drosophila* and *Xenopus* embryos revealed that inactivation of *Zw3/GSK-3β* results in the accumulation of the protein β-catenin/armadillo (Peifer *et al.*, 1994; Hinck *et al.*, 1994; Yost *et al.*, 1996). GSK-3β phosphorylates β-catenin at a conserved N-terminal site and mutations within β-catenin at these sites of phosphorylation results in a more stable form of β-catenin (Rubinfeld *et al.*, 1996). GSK-3β also phosphorylates APC and numerous studies have shown that site specific phosphorylation of APC within the kinase regulated β-catenin binding domain is crucial for the binding of β-catenin to this domain and subsequently degradation of β-catenin (2.3.3 and 2.6.7.2). Axin is also phosphorylated by GSK-3β, discussed in 2.6.7.3.

GSK-3β, Axin, β-catenin and APC are thought to form a tetrameric complex allowing efficient phosphorylation of these GSK-3β targets mediating degradation of β-catenin. Axin can be viewed as the scaffolding protein to these interactions as all components bind to this protein directly, depicted in figure 8.

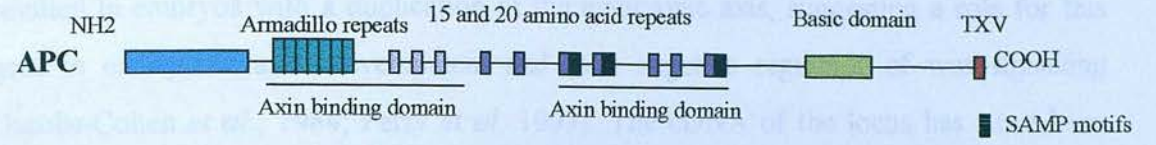
Figure 8: Axin is the scaffolding protein in a protein complex that regulates β-catenin stability.

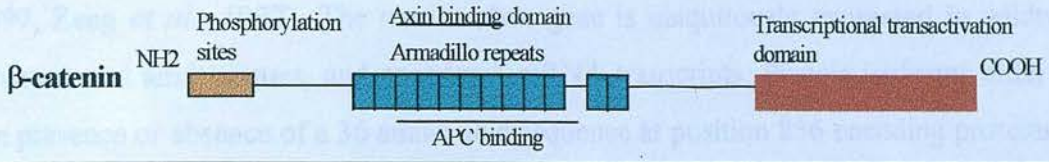
Figure 8a depicts protein-binding domains identified within the APC, β-catenin and Dsh proteins. Figure 8b represents the protein complex that is thought to form to down regulate the level of cytoplasmic β-catenin.

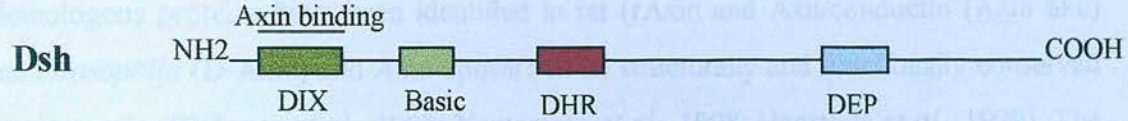
The binding domains and protein interactions are fully described within the text of section 2.6.

Figure 8: Axin is the scaffolding protein in a protein complex that regulates β-catenin stability

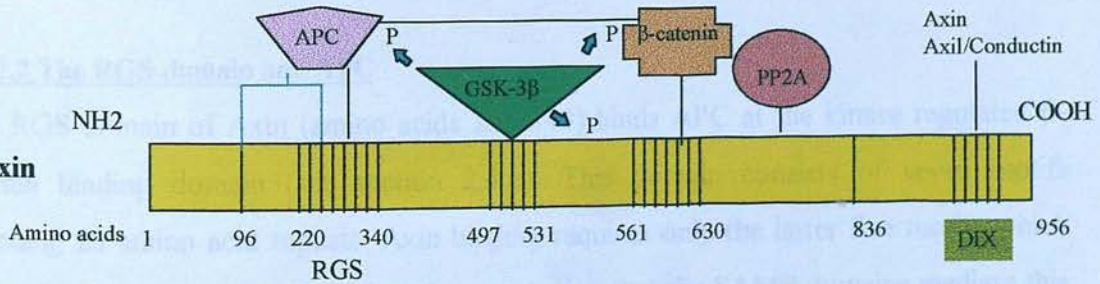
A: Protein interaction domains







B: Protein complex



2.6.7 GSK-3β/Axin/β-catenin and APC

2.6.7.1 Axin

Axin was originally identified as a product of the mouse *Fused* gene (Zeng *et al.*, 1997). Some spontaneously occurring alleles of the Fused gene, when exhibited homozygously, resulted in embryos with a duplication of the embryonic axis, suggesting a role for this gene in embryonic axial development and as a negative regulator of wnt signalling (Jacobs-Cohen *et al.*, 1984; Perry *et al*, 1995). The cDNA of the locus has since been sequenced and the *Fused* gene was renamed *Axin* (Perry *et al.*, 1995; Vasicek *et al.*, 1997; Zeng *et al.*, 1997). The murine *Axin* gene is ubiquitously expressed in wildtype embryos and adult tissues, and encodes 3 mRNA transcripts. Protein isoforms differ by the presence or absence of a 36 amino acid sequence at position 856 encoding proteins of either 956 or 992 amino acids.

Homologous proteins have been identified in rat (rAxin and Axil/conductin (Axin like) and *Drosophila* (D-Axin) and Axin appears to be structurally and functionally conserved across species (Behrens *et al.*, 1998; Yamamoto *et al.*, 1998; Hamanda *et al.*, 1999). The protein structure includes an N-terminal region which shows homology to the regulator of G protein signalling (RGS) domain, the central domain contains sites shown to interact with β-catenin, GSK-3β and protein phosphatase 2A, and the carboxyl terminal contains a DIX domain.

2.6.7.2 The RGS domain and APC

The RGS domain of Axin (amino acids 220-340) binds APC at the kinase regulated β-catenin binding domain (see section 2.3.3). This domain consists of seven motifs encoding 20 amino acid repeats. Axin binding requires only the latter five motifs which contain the interspersed SAMP domains suggesting that the SAMP domains mediate this interaction (Behrens *et al.*, 1998; Ikeda *et al.*, 1998; Kishida *et al.*, 1998; Hart *et al.*,

1998; Fagotto *et al.*, 1999).

A second APC binding domain within Axin has been recently identified, the region between amino acids 96-253 can directly bind to the NH2-terminal region of APC which encodes the Armadillo motifs and the constitutive β-catenin binding domain (Fagotto *et al.*, 1999). In the absence of the RGS domain APC/Axin interactions were not detected suggesting that this may be a secondary binding site for already bound APC. It is unclear whether Axin and β-catenin compete for APC binding.

2.6.7.3 Axin and GSK3β

Several immunoprecipitation studies have identified direct interactions between Axin and GSK-3β, although experiments have failed to detect the interaction of *Drosophila* homologues, zw-3 and D-Axin (Hamanda *et al.*, 1999). The region of Axin required for GSK-3β binding is located between amino acids 497-552 (Ikeda *et al.*, 1998; Kishida *et al.*, 1998; Fagotto *et al.*, 1999; Smalley *et al.*, 1999). *In vitro* studies elucidated that catalytically inactive mutants of GSK-3β are unable to interact with Axin suggesting that prior phosphorylation of Axin is required for the interaction of these two proteins (Kishida *et al.*, 1998). Furthermore, *in vivo* experiments have shown that binding of GSK-3β to Axin is required for GSK-3β mediated phosphorylation of APC (*Fagotto et al.*, 1999).

2.6.7.4 Axin and β-catenin

In vitro studies have determined that Axin and its homologues Axil/conductin can interact with the Armadillo repeat domain of β-catenin, and this interaction is mediated through a highly conserved region present in Axin and its homologues (Zeng *et al.*, 1997; Behrens *et al.*, 1998; Ikeda *et al.*, 1998; Sakanaka *et al.*, 1998; Yamamoto *et al.*, 1998; Kishida *et al.*, 1998; Itoh *et al.*, 1998; Hamanda *et al.*, 1999). Interactions between β-catenin and either rAxin or Axil promotes GSK-3β dependent phosphorylation of β-catenin (Ikeda *et al.*, 1998; Yamamota *et al.*, 1998).

The Armadillo repeat domain of β-catenin is also the site of interaction with APC, α-

catenin, cadherin and the LEF-1/TCF family of proteins (Hülsken *et al.*,1994a; Funayama *et al.*, 1995; Behrens *et al.*, 1996; Molenaar *et al.*, 1996; Polakis *et al.*, 1997). APC binds β-catenin between armadillo motifs 2 and 10 whilst rAxin binds β-catenin within the region of motifs 2 and 7. As previously discussed in section 2.3.2, armadillo motifs are not 100% homologous, therefore different motifs or groups of motifs may represent different sites of protein interaction. As Axin has distinct binding sites for APC (RGS domain) and β-catenin (β-catenin binding domain) and APC has distinct binding sites for Axin and β-catenin these proteins are thought to form a ternary complex. Further evidence that these components exist in a ternary complex comes from immunoprecipitation studies using site specific mutants (Ikeda *et al.*, 1998; Kishida *et al.*, 1998; Fagotto *et al.*, 1999).

2.6.7.5 Axin, phosphatase 2A and the carboxyl terminal domain

Amino acids 632-836 of the Axin protein bind to the catalytic subunit of the Ser/Thr protein phosphatase 2A (PP2A) (Hsu *et al.*, 1999). The binding of PP2A to Axin may suggest that PP2A could positively regulate wnt signalling by opposing the action of GSK-3β. Studies have demonstrated that unphosphorylated Axin is less stable than when in a phosphorylated form (Yamamoto *et al.*, 1999). One could hypothesise that PP2A could promote destabilisation of Axin and the Axin based protein complex resulting in increased β-catenin levels and wnt signalling or equally, modification of APC and/or β-catenin phosphorylation could result in the same effect.

A role for the PP2A binding in the negative regulation of wnt signalling is supported by the phenotype seen in mice that harbour the spontaneous Fused alleles (see section 2.6.7.1). *Fused* (*Axin*^{FU}) and *Knobbly* (*Axin*^{KB}) encode proteins with C-terminal truncations that are predicted to retain binding domains for APC, GSK-3β and β-catenin, but reduce or abolish the PP2A binding site. The phenotype seen in these mice is typical of overexpression of the wnt signalling pathway.

Analysis of the C-terminal domain of Axin revealed a self binding domain located within the terminal 100 amino acids of the protein (Hsu *et al.*, 1999; Sakanaka and Williams

1999). This fragment contains the DIX domain, a 51 amino acid region of similarity between Axin and Dsh proteins (Hsu *et al.*, 1999). Therefore, Axin may have the ability to homo- or hetero-dimerise with different Axin isoforms or even the closely related protein Axil/conductin (Behrens *et al.*, 1998; Yamamoto *et al.*, 1998). It is unclear what function Axin homodimerisation could play, but possibly dimerised molecules could further regulate β-catenin signalling, enhancing interactions between proteins within the complex or in some way regulate phosphorylation.

Dsh has been shown to interact with Axin through the Dsh DIX domain, see section 2.6.5 (Kishida *et al.*, 1999; Smalley *et al.*, 1999), however, Smalley and co-workers demonstrated that this interaction does not require the DIX domain of Axin.

2.6.8 β-catenin degradation via the ubiquitin-proteosome pathway

Studies have shown that β-catenin is a target for the ubiquitin-proteosome pathway and that phosphorylation of serine/threonine residues at positions 29,33,37,41 and 45 by GSK-3β appear to be a prerequisite for ubiquitination (Aberle *et al.*1997; Orford *et al.*, 1997). The ubiquitin proteosome pathway plays a critical role in diverse biological processes though the rapid elimination of many important regulatory proteins. Ubiquitin mediated proteolysis (reviewed in Hochstrasser, 1995) involves a cascade of ubiquitin transfer reactions in which the ubiquitin-activating enzyme E1 uses ATP to form a high energy thiolester bond with ubiquitin which is then transferred to members of the E2 ubiquitin-conjugating enzyme family. Ubiquitin is then transferred to lysine residues in the target though E3-ubiquitin ligases, which can be either a single polypeptide, or, in many cases, a protein complex. E3 serves as an adapter that interacts with both the target protein and the appropriate E2 enzyme thereby providing specificity to the ubiquitin transfer reaction. Multiple rounds of ubiquitination lead to poly-ubiquitination, which targets the protein for proteolysis by the 26S proteosome.

2.6.9 β-catenin LEF/TCF

Wnt signalling inhibits β-catenin degradation, and β-catenin accumulates within the cytoplasm. β-catenin is subsequently translocated to the nucleus (reviewed in Kuhl and Wedlich 1997). β-catenin does not encode a recognised nuclear localisation sequence (NLS) but can be imported into the nucleus by binding directly to the nuclear pore machinery (Fagotto *et al.*, 1998). In the nucleus β-catenin associates with lef-1/tcf family members.

Tcf-1 and *lef-1*, the two founding members of the small subfamily of vertebrate high-mobility group (HMG) box transcription factor genes were originally defined as lymphoid-specific transcription factors. This was based on their affinity for the enhancers of CD3ε and the T-cell receptor α genes respectively. Both genes were later found to be expressed in largely overlapping complex patterns during mouse embryogenesis (Sawada and Littman 1991, Travis *et al.*, 1991; Waterman *et al.*, 1991; van de Wetering *et al.*, 1992, Oosterwegel *et al.*, 1993; Carlsson *et al.*, 1993). To date four members of this family have been identified in mammals: lef-*1*, tcf-*1*, tcf-*3* and tcf-*4* (Travis *et al.*, 1991; Waterman *et al.*, 1991; van de Wetering *et al.*, 1992; Carlsson *et al.*, 1993; Morin *et al.*, 1997), one member in *Drosophila*: *D-tcf/Pangolin* (van de Wetering *et al.*, 1997) and one member in *C.elegans*; *POP-1* (Lin *et al.*, 1995).

All proteins contain a highly homologous β-catenin binding domain located at the amino terminal, a carboxyl HMG DNA binding domain and a central contex dependent activation domain (CAD) that is divergent between members. The β-catenin binding domain of Lef-1/tcf interacts with β-catenins armadillo repeat domain, and armadillo repeats 1-7 have been shown to be required for this interaction (Behrens *et al.*, 1996; Hsu *et al.*, 1998). The HMG DNA binding domains, comprise three α-helices arranged in an L-shape which binds to highly similar AT-rich DNA motifs or a more conserved 11 bp motif encoded by promoters of target genes (van wetering *et al.*, 1997). Transcriptional activation function is mediated by the CAD, although the mechanism of transcriptional transactivation is unclear. Within the lef-1/tcf and β-catenin complex, the

amino terminal 131 residues and the carboxyl terminal region of β-catenin is thought to confer transcriptional transactivation.

Analysis of dominant negative forms of *Lef*-1/*Tcf*, that lack the β-catenin binding domain, block signalling mediated by β-catenin or wnt (Molenaar *et al.*, 1996; van de Wetting, 1997; Kengaku *et al.*, 1998; Dorsky *et al.*, 1998) and loss of function mutants of *D-tcf* in *Drosophila* have shown that D-tcf is required for downstream effects of wnt signalling (van de Wetering, 1997; Brunner *et al.*, 1997). Analysis of *Lef*-1/*Tcf* mutations in the mouse has resulted in diverse phenotypes. Targeted mutations of the *Lef-1* gene results in developmental abnormalities of multiple organs including hair follicles, the mammary glands and teeth (van Genderen *et al.*, 1994), whilst mutations in *Tcf-1* impairs T-cell differentiation (Verbeek *et al.*, 1995). Mice null for *Tcf-4* have defects in colonic stem cell proliferation (Korinek *et al.*, 1998). These phenotypes suggest that specific Lef-1/Tcf members are critical in different tissues/cell types, although it is possible there is functional redundancy between Lef-1/Tcf transcription factors and this is supported by the overlapping expression pattern detected in mouse embryos (Oosterwegel *et al.*, 1993; Korinek *et al.*, 1998). Furthermore, recently it has been documented that mice null for both *Lef-1* and *Tcf-1* exhibit defects in mesoderm specification that are virtually identical to those observed in *Wnt3a* deficient mice (Galceran *et al.*, 1999).

In the absence of wnt signalling and stabilisation of β-catenin, lef-1/tcf proteins actively repress transcription of wnt responsive genes. *Ultrabithorax* (*ubx*) and *Siamosis* are genes that are responsive to wnt signals in *Drosophila* and *Xenopus* respectively. Mutations within the lef-1/tcf binding region of these genes resulted in ectopic gene expression, this lead to the proposal that lef-1/Ttcf proteins act as repressors (Riese *et al.*, 1997; Brannan *et al.*, 1997). Further to this, lef-1/tcf proteins have been shown to interact with the transcription corepressor Groucho and CBP (CREB binding protein) (Roose *et al.*, 1998; Cavallo *et al.*, 1998; Waltzer and Bienz, 1998).

2.6.9.1 Transcription corepressors

Promoter activity in eukaryotic cells is regulated to a certain degree by coactivators and

corepressors (reviewed in Kadonaga *et al.*, 1998). These factors modulate rates of transcription by a variety of mechanisms, they do not have the ability to bind DNA on their own, but are recruited to target promoters by protein-protein interactions with DNA-binding transcription factors.

Groucho is the founding member of a family of transcriptional corepressor proteins that contain a highly conserved N-terminal region (the effector of repression) and multiple tandemly repeated copies of a 40 amino-acid motif known as the WD40 repeat (Reviewed in Parkhurst, 1998; Fisher and Caudy, 1998). The WD40 repeat domain interacts with WRPW (Tryptophan-Arginine-Proline-Tryptophan) and WRPY (Y-Tryrosine) motifs that are present in a number of transcription factors.

Evidence suggests that the interaction of lef-1 or tcf with Groucho results in repression of lef-1/tcf target genes. When β-catenin enters the nucleus it forms a complex with lef-1/tcf transcription factors and target genes are transactivated. It could be hypothesised that β-catenin displaces Groucho and relieves repression of transcription.

CBP, a histone acetyltransferase (reviewed in Imhof *et al.*, 1997), is a modulator of tcf/β-catenin binding. *In vitro*, CBP has been shown to bind the HMG domain of D-tcf, which results in the acetylation of lysine residues in the β-catenin binding region of tcf. Acetylation lowers the affinity of the interaction between β-catenin and tcf, this modification is thought to inhibit signalling until the levels of Armadillo reaches a threshold (Waltzer and Bienz, 1998).

2.6.10 Responsive genes of Wnt/β-catenin signalling

Based on genetic evidence, overexpression studies and promoter sequence analysis, target genes have been identified in *Drosophila*; *Ultrabithorax* (Riese *et al.*, 1997; van de wetering *et al.*, 1997); *Xenopus*: *Siamosis*, and *nodal-related-3* (Brannon *et al.*, 1997; Mckendry *et al.*, 1997) and mammals: *connexin 43*: *CCND1*, *c-MYC* and *matrilysin* (van der Heyden *et al.*, 1998; Tetsu *et al.*, 1999; Shtutman *et al.*, 1999; He *et al.*, 1998). The identification of these genes now provides a link with the diverse biological processes

associated with the wnt signalling pathway.

2.6.10.1 *Ultrabithorax*

Ubx is a member of the family of homeobox transcription factors expressed in *Drosophila. Ubx* functions to signal endoderm induction and cell type specification during gut development (Bienz and Tremml, 1988; Bienz, 1994). Reciprocal inductive interactions between different tissues that result in differentiation and organisation of cells into organs are crucial for morphogenesis of vertebrates and invertebrates. Analysis of the *Ubx* promoter sequence revealed a lef-1/tcf binding motif and investigations have demonstrated that lef-1 can activate *ubx* transcription (Riese *et al.*, 1997). Mice deficient for *Lef-1* exhibit defects in organs that require inductive epithelial-mesenchymal interactions during organogenesis, specifically mutant mice lack teeth, mammary glands, whiskers and body hair (van Genderen *et al.*, 1994).

2.6.10.2 *Siamosis*

Siamosis is a homeobox gene that is expressed in the dorsal vegetal cells soon after the start of zygotic transcription. Ectopic expression of *siamosis* induces complete axis duplication in xenopus (Lemaire *et al.*, 1995) The *siamosis* promoter has been shown to be a direct target of the β-catenin/ X-tcf-3 complex and multiple binding sites for X-tcf-3 have been identified within the promoter (Brannon *et al.*, 1997). Furthermore, investigations have demonstrated that X-tcf-3 is a repressor of the *siamosis* promoter (Brannon *et al.*, 1997).

2.6.10.3 nodal-related 3

Nodal-related 3 (*Xnr3*) is a key regulator of *Xenopus* development (Smith *et al.*, 1995; McKendry *et al.*, 1997). *Xnr3* encodes a protein related to members of the TGFβ family of signalling molecules and is specifically expressed in the epithelial layer of the spemann organiser (a broad zone on the dorsal side of the embryo) during gastrulation (Smith *et al.*, 1995). Xnr3 has both inductive and morphogenic activity in the gastrula embryo

(Hansen *et al.*, 1997). The *Xnr3* promoter was shown to confer wnt inducible transcription of a reporter gene. Two distinct regulatory sequences were identified, both mediated wnt responsiveness. One regulatory element was identified as the binding site of the LEF-1/TCF family of transcription factors, whereas the other sequence element interacted with an unknown protein from *Xenopus* gastrulae.

2.6.10.4 *Connexin 43*

Connexin 43 (Cx43) is one of a family of proteins controlling the permeability of gap junctions (reviewed in Goodenough *et al.*, 1996). Gap junctions are intercellular channels that permit exchange of ions, small metabolites and low molecular mass signalling molecules (Goodenough *et al.*, 1996). *Cx 43* was identified as a target gene of wnt signalling following overexpression of *Wnt1* in a rat neuronal cell line (PC-12) (Bradley *et al.*, 1995). Overexpression of *Wnt1* resulted in a tendency for cells to grow in highly connected groups, indicating that wnt signalling can result in increased cellular adhesion. Further studies revealed that this phenotype was a result of increased coupling by gap-junctions, measured by enhanced chemical and electrical gap-junction communication. Subsequent experiments revealed elevated levels of Cx43 (van der Heyden *et al.*, 1998), and this phenotype was also noted following experiments on a mammary epithelial cell line. Multiple lef-1/tcf consensus binding sequences are present in rat; human and murine CX43 promoters.

Van der Heyden and co-workers (1998) also demonstrated that treatment of cells with lithium, an inhibitor of GSK-3β kinase activity, resulted in increased transcription from the *CX43* promoter, further associating Cx43 with wnt signalling pathway. It is interesting to note that mice deficient in *Wnt-1* or which overexpress *Cx43* both exhibit severe brain defects (Thomas and Capecchi 1990; Nusse and Varmus 1992; Ewart *et al.*, 1997).

2.6.10.5 *c-MYC*

c-MYC plays a role in both positive and negative growth through its influence on proliferation, differentiation and apoptosis (Reviewed in Hoffman and Liebermann 1998; discussed in detail in section 1.3.1.2). Using serial analysis of gene expression (SAGE), a technique that allows the quantitative evaluation of cellular mRNA (Velculescu *et al.*, 1995), investigators evaluated gene expression in the colon carcinoma cell line HT29, a cell line that contains truncating mutations within the *APC* gene, following exogenous expression of an inducible *APC* construct. Several transcripts were shown to be overexpressed (14) and repressed (16) by *APC* induction. As APC represses β-catenin/TCF-4 mediated transcription (Korinek *et al.*, 1997), repressed transcripts were analysed further. One highly repressed transcript was identified as c-MYC mRNA. This repression was confirmed at the mRNA and protein level by northern blot and immunoblot analysis respectively. In the human colon cancer cell line SW480 (a cell line also encoding only truncated APC), expression of the inducible *APC* construct was shown to significantly repress the transcriptional activity of a *c-MYC* reporter plasmid (*c-MYC* promoter inserted upstream of a luciferase reporter gene). Conversely, expression of a mutant form of β-catenin, that rendered the protein insensitive to GSK-3β phosphorylation and subsequent degradation, when expressed in the human kidney cell line 293, resulted in significant activation of the *c-MYC* reporter plasmid.

Analysis of the *c-MYC* promoter sequence revealed two putative TCF binding sites. Mutations within either of these sites reduced activity of the *c-MYC* promoter by 50%. Deletion of both sites removed *APC* repression and *β-catenin* activation of the *c-MYC* reporter plasmid.

The exact mechanisms by which c-MYC mediates cell proliferation, apoptosis and differentiation are still unclear but numerous investigations both *in vivo* and *in vitro* have demonstrated that deregulation of *c-MYC* can result in tumourigenesis. *C-MYC* is overexpressed due to genetic alterations including chromosomal translocations, proviral insertion, retroviral transduction and gene amplification (Ryan and Birnie, 1997). However, *c-MYC* is overexpressed in colorectal cancers and NSCLC cancers that do not

exhibit gene amplification. In the cases of colorectal cancers, where overexpression of *c-MYC* is reported to occur in approximately 70% of cases (Sikora *et al.*, 1987), mutations within the *APC* or the *β-catenin* genes occur in the majority of cases (Sparks *et al.*, 1998). The identification of *c-MYC* as a target gene of wnt signalling could therefore explain aberrant expression.

In NSCLC, LOH at 5q21, the location of the *APC* gene, is relatively common although studies have not fully addressed the possibility of biallelic inactivation of the *APC* gene (see section 1.3.3.3). No reports of mutations in other components of the wnt signalling pathway have been reported. As previously discussed, activation of the proto-oncogene *K-RAS* occurs in approximately one third of NSCLC cases, as *c-MYC* is also one of the target genes of the MEK/ERK signalling pathway then this may also be a mechanism by which *c-MYC* is deregulated (see section 1.3.1.1).

2.6.10.6 Cyclin D1

Cyclin D1 is a major regulator of the progression of cells into the proliferative stage of the cell cycle (reviewed in Sherr 1996; discussed in section 1.3.1.4) and has recently been identified as a target of wnt signalling (Tetsu and McCormick 1999; Shtutman *et al.*, 1999). Four LEF/TCF consensus sequences have been identified in the promoter region of the *CCND1* gene. All sites were shown to confer cyclin D1 activation *in vitro* and inactivation of increasing numbers of sites resulted in a gradual reduction in responsiveness to β-catenin.

Overexpression of cyclin D1 has been reported to occur in approximately 50% of NSCLC cases, of which amplification of the *CCND1* accounts for around 2-30% cases (see section 1.3.1.4). Oncogenic activation of *K-RAS* resulting in transcriptional activation of cyclin D1 via the MEK/ERK and PI3K pathways is thought to contribute to cyclin D1 expression seen in a high proportion of these cases. Alternatively, deregulation of cyclin D may be a result of loss of *RB* expression (see section 1.3.2.1). The identification of Cyclin D1 as a target gene in the wnt signalling pathway now adds another pathway to be investigated.

2.6.10.7 Matrilysin

Greater than 90% of colonic adenomas that harbour germline mutations in the *APC* gene (in both human and mice) have been shown to overexpress matrilysin (Wilson *et al.*, 1997; Fingleton *et al.*, 1999). Matrilysin was therefore considered a possible target gene of β-catenin/tcf transactivation. Crawford and co-workers (1999) confirmed this in a series of experiments. In murine intestinal adenomas and ADC, matrilysin transcripts were shown to co-localise with areas of β-catenin protein accumulation. The mouse matrilysin promoter was cloned and sequenced and revealed a single canonical Tcf-4 binding site. Using electrophoretic mobility shift assays (EMSA), the *Tcf* motif was shown to interact with *Tcf-4* and the β-catenin/Tcf bipartite complex, these interactions were eliminated by the introduction of a 2 base pair mutation within the *Tcf* binding motif. Further *in vitro* experiments demonstrated that the *Tcf-4* binding site exhibited transcriptional repressor properties and that formation of the β-catenin/Tcf complex relieves gene repression mediated by Tcf only, but does not directly act as a transcriptional activator of the Tcf binding site (Crawford *et al.*, 1999).

The matrix metalloproteinases (MMPs) are a family of secreted, zinc-dependent enzymes known to degrade components of the extracellular matrix (Reviewed in Matrisian, 1992). Members of the metalloproteinase family can be divided into three categories depending on substrate specificity; the collagenases, gelatinases and stromelysins. Matrilysin is a member of the stromelysin category that has been shown to degrade proteoglycans, gelatins, elastin and glycoproteins such as fibronectin, laminen and entactin (Matrisian 1992). The regulation of MMPs is complex with selective expression in specific cell types being induced by a variety of growth factors, cytokines and hormones. MMP expression can alter the integrity of the extracellular matrix (ECM) surrounding the cell and in turn affect cellular morphology and cellular processes such as proliferation, migration or differentiation (reviewed in Matrisian 1992). *In vivo* experiments have demonstrated that expression of MMPs can induce tumour phenotypes in well-differentiated cells (Witty *et al.*, 1994). Virtually all members of the MMP family have been shown to be

overexpressed in malignant neoplasm of various origins, in most cases overexpression has been associated with the invasive and metastatic stages of tumour progression (Matrisian, 1992). MMPs are thought to facilitate tumour invasion and metastasis by breaking down membrane and underlying stroma, however, evidence suggests that matrilysin overexpression is not sufficient for tumour invasion and metastasis, and other factors are involved (Lochter *et al.*, 1997).

Matrilysin is distinct from other MMPs in its expression pattern, localising to the epithelial component of ADC (Pajouh *et al.*, 1991; McDonnell *et al.*, 1991; Newell *et al.*, 1994) whereas other MMP transcripts generally localise to the stromal components (Muller *et al.*, 1993; Newell *et al.*, 1994; reviewed in Powell and Matrisian 1996). A role for matrilysin in colorectal carcinogenesis has been demonstrated through the analysis of Min (multiple intestinal neoplasia) mice. Min mice harbour a germline mutation in *Apc* gene that results in the FAP phenotype. When Min mice were crossed to matrilysin deficient mice, fewer and smaller colorectal adenomas were detected (Wilson *et al.*, 1997).

2.7 APC and cell adhesion

Adhesive interactions between cells and the interaction of cells with the extracellular matrix play a critical role in a number of biological activities including embryonic and organ development and maintenance of tissue integrity (Takeichi 1995; Gumbiner 1996; Vleminckx 1999). Cell adhesion is mediated by a number of different molecules that include members of the integrin, cadherin, immunoglobulin and proteoglycan protein families. Integrins and proteoglycans can execute cell-matrix, as well as cell-cell interactions whilst adhesion via cadherins appears to be strictly intercellular. Cadherins are of particular interest as they complex with β-catenin. Thus β-catenin functions in both wnt signalling and cell adhesion.

The term cadherin denotes calcium-dependent intercellular adhesion molecules. Classical cadherins are defined by their characteristic and highly conserved cytoplasmic domain and include E- (epithelial), N- (neuronal) and P- (placental) cadherin which are primarily found at the adherens junctions of adjacent cells (Kemler 1992; Yap *et al.*, 1997). Cadherins exist at the cellular membrane as parallel dimers consisting of the same family members. The extracellular segment consists of 5 cadherin repeats, EC1-EC5; these repeats contain calcium binding sub-domains. Calcium is critical for cadherin function and serves to maintain the structural integrity of the transmembrane protein. Different family members express various EC-1 subunits and cells adhere preferentially to cells expressing the identical family member. The highly conserved cytoplasmic tail of the classical cadherins serves as a link to the cytoskeleton of the cell through specific interactions with the intracellular proteins α-, β- and γ- catenin (Gumbiner 1993; Ranscht 1994). β- and γ catenin share a high degree of homology and interact with the cytoplasmic tail of cadherin through armadillo repeat sequences located in the central portion of the proteins (Hülsken *et al.*, 1994a; Funayama *et al.*, 1995). β-catenin and γ-catenin interact with α-catenin linking the cadherin-catenin complex, either directly or indirectly ia the actin binding proteins α-actin or vinculin, to the actin filament network (Rimm *et al.*, 1995; Knudsen *et*

al., 1995; Weiss *et al.*, 1998). This interaction is essential for cell-cell adhesion, and its regulation is critical for the highly dynamic nature of cellular junctions during for example, embryonic development or wound healing where junctions have to be rapidly disassembled and reassembled (Angres *et al.*, 1996; Marrs and Nelson 1996).

Another catenin that associates with the cadherin adhesion complex is $p120^{ctn}$ (Reynolds *et al.*, 1994; Aghib and McCrea 1995; Shibamoto *et al.*, 1995; Staddon *et al.*, 1995). $P120^{ctn}$ appears to have a distinctive biological role from β-catenin and γ-catenin. $P120^{ctn}$ binds to a more proximal region of the cadherin cytoplasmic tail than β- and γ-catenin (Yap *et al.*, 1998). Data implicate a role in cell mobility (Riehl *et al.*, 1996; Chen *et al.*, 1997) and in cadherin clustering and adhesive strengthening (Yap *et al.*, 1998).

Besides physically linking cadherin to the cytoskeleton, the cytoplasmic domain of cadherins can interact, either directly or indirectly (through the catenins) with other proteins that have potential signal transducing activities. These interacting proteins include the receptor tyrosine kinases erb-b1 and erb-b2, the cytoplasmic kinases (such as SRC) and the adaptor protein SHC (Daniel and Reynolds 1997; Xu *et al.*, 1997). How cadherins modulate cellular signalling is generally unclear but this signalling in addition to the mechanism of cell adhesion, are thought to mediate cellular responses associated with cadherin proteins such as cell proliferation, apoptosis and cell differentiation (reviewed in Vleminckx and Kemler, 1999).

2.7.1 Interaction of β-catenin with cadherins and the wnt signalling pathway

β-catenin plays a crucial role in both cadherin-mediated cell-cell adhesion and wnt signal transduction. Three proteins that facilitate these functions are APC, Lef-1 and cadherin, which mediate β-catenin degradation, activation of gene transcription and cell adhesion respectively. All three proteins bind to the armadillo repeat region of β-catenin. The binding sites of these proteins have been shown to partially overlap and experiments have demonstrated that proteins compete for mutually exclusive binding of β-catenin (Hülsken *et al.*, 1994b; Rubinfeld *et al.*, 1995; Behrens *et al.*, 1996; Huber *et al.*, 1996; Orsulic and

Peifer, 1996; Orsulic *et al.*, 1999).

In a simple cellular model, β-catenin is sequestered and stabilised at the cell membrane, non-sequestered or free β-catenin can bind APC and forms a complex with Axin, this complex is phosphorylated by GSK-3β and β-catenin is targeted for degradation. In the presence of wnt signals, GSK-3β activity is inhibited and β-catenin levels rise saturating cadherin molecules, excess β-catenin then translocates to the nucleus and activates target genes.

Several *in vivo* studies have demonstrated that cadherins can antagonise wnt signalling by competing for β-catenin binding (Sanson *et al.*, 1996; Fagotto *et al.*, 1996; Torres *et al.*, 1996). Recently, Orsulic and co-workers (1999) have also demonstrated that loss of E-cadherin, an event that is commonly associated with the progression of the majority of epithelial tumours, mimics wnt signalling (Behrens *et al.*, 1989; Vleminckx *et al.*, 1991; Birchmeier and Behrens, 1994). Therefore, it could be hypothesised that loss of E-cadherin could promote tumourigenesis not only as a result of deregulated cellular adhesion, but also by the promotion of transcription of wnt responsive genes such as *CCND1*, *c-MYC* and matrilysin. Alternatively, one could consider deregulation of β-catenin levels, accumulation of β-catenin within the cytoplasm may lead to increased cellular adhesion in cases where E-cadherin expression is not lost, i.e. in the early stages of tumourigenesis.

Chapter 3: Loss of heterozygosity at chromosome 5q21 and gene mutation in the *APC* gene in NSCLC

3.1 Introduction

The functional inactivation of tumour suppressor genes is a common event in neoplasms. According to Knudson's hypothesis, tumour suppressor genes act recessively at the cellular level so that both copies of the gene must be inactivated for growth suppressive functions to be eliminated (Knudson 1978). In hereditary syndromes, generally a germline mutation in one allele predisposes an individual, and following inactivation of the remaining wildtype allele (by mutation or allele loss through partial arm deletion or mitotic recombination) gives rise to a phenotypic effect. In sporadic neoplasms, inactivation of a tumour suppressor gene requires two genetic events to occur (Weinberg, 1991). Previous chapters have discussed several tumour genes that are inactivated in NSCLC.

The chromosomal region 5q21 is deleted in numerous neoplasms including NSCLC where results have shown loss in up to 70% of cases (Ashton-Rickardt *et al.*, 1991; D'Amico *et al.*, 1992; Tsuchiya *et al.*, 1992; Horri *et al.*, 1992a; Hosoe *et al.*, 1994; Wieland *et al.*, 1994; Fong *et al.*, 1995a). Within NSCLC, a common region of chromosomal deletion has been identified (Wieland *et al.*, 1994). This region spans 3-5 megabases and harbours numerous genes including the tumour suppressor gene *APC*. Germline mutations within the *APC* gene predisposes individuals to FAP, a disease whereby multiple adenomas develop within the colon predisposing the individual to colorectal cancer and in some instances extracolonic manifestations (see section 2.4). Analysis of subsequent tumours reveals mutation or loss of the remaining *APC* allele. Biallelic inactivation of the *APC* gene has also been reported in sporadic colorectal cancer cases (Nishisho *et al.*, 1991; Groden *et al.*, 1991) and gastric and hepatic cancers (Tamura *et al.*, 1994; Oda *et al.*, 1996; Imai *et al.*, 1997).

Gene mutations have also been detected in pancreatic (Yashima *et al.*, 1994; Horri *et al.*,

1992a), oesophageal (Powell *et al.*, 1994; Gonzalez *et al.*, 1997) and prostate cancers (Watanabe *et al.*, 1996). In these neoplasms, only one *APC* allele has been shown to be inactivated and it is unclear whether APC has a role in promoting tumourigenesis. If Knudson's hypothesis were to be up held then inactivation of one allele would not result in a phenotypic change. However, in the cases were the allele has been subject to a mutation, rather than loss of the whole gene, then one must consider the possibility of a dominant negative effect (see sections 2.3.1 and 2.4).

Previous publications report that LOH within the chromosomal region 5q21 occurs in 20-70% of NSCLC cases. Several factors could be responsible for the variability in frequency between cohorts. Firstly, some studies included NSCLC cell lines, the results from these studies can be unreliable as often during culturing cell lines acquire genomic aberrations. Secondly, the relative percentages of histological subtypes (SCC, ADC and LCC) vary. A previous investigation reported that ADC and SCC exhibit different allelotypes (Sato *et al.*, 1994) suggesting that different genes may be involved in driving tumourigenesis in these different histological groups. Two studies have analysed the frequency of LOH at 5q21 in ADC and SCC. Results were conflicting with one group reporting that LOH at 5q21 is more frequent in ADC than SCC whilst the other reported LOH to be more frequent in SCC (Tsuchiya *et al.*, 1992; Fong *et al.*, 1995a). Conflicting results between these studies may be due to the frequency of different grade tumours as Fong and co-workers (1995a) showed that LOH at 5q21 was more common in advanced stage NSCLC suggesting that LOH at 5q21 is a late event in NSCLC tumourigenesis.

3.2 Aims

In the following study, my aim was to determine whether the *APC* tumour suppressor gene has a role in the tumourigenesis of NSCLC. Initial investigations analysed the frequency of LOH at 5q21 in a cohort of well-characterised NSCLC cases. This was carried out through the analysis of polymorphic sites within the *APC* and *MCC* genes.

Secondly, I analysed the frequency of LOH at 5q21 in ADC and SCC subtypes. LOH at 5q21 was also analysed following the subclassification of ADC into parenchymal and bronchial origin, as described by Edwards (1987), discussed in section 1.2.1.2.

Thirdly, the frequency of LOH at 5q21 was analysed with respect to tumour grade.

These initial studies address LOH at 5q21 and analyse results with respect to histological subtype and tumour grade. The final aim of this investigation was to determine whether biallelic inactivation of the *APC* gene is a characteristic of NSCLC. NSCLC cases shown to exhibit LOH at 5q21 were subject to mutational analysis of the *APC* gene. Two methods of analysis were chosen. The first method was single stranded conformation polymorphism analysis (SSCP). This is a relatively quick method to screen fragments of DNA for mutations. As the majority of mutations, both sporadic and germline, have been shown to occur within the mutation cluster region (MCR) of the *APC* gene, SSCP was used to analyse this region. As SSCP analysis is not 100% reliable and because mutations may also occur outwith the MCR, NSCLC cases were also subject to immunohistochemical analysis. An antibody raised to the carboxyl terminal region of the APC protein was used in this technique as the majority of mutations result in the production of a truncated protein product.

3.3 Materials and Methods

3.3.1 Tissue collection

A retrospective analysis of paraffin embedded non-small cell lung cancer (NSCLC) biopsies was carried out. A total of 99 surgically resected lung cancers were analysed, consisting of 33 consecutive squamous carcinomas and 59 consecutive adenocarcinomas and seven additional bronchial adenocarcinomas. The additonal bronchial adenocarcinmas were included to improve statistical analysis between parenchymal and bronchial adenocarcinomas as bronchial adenocarcinomas occur less frequently (Dr. F. Carey, personal communication, Edwards 1987). For each NSCLC case, consecutive sections of 3 and 10μm were cut using a microtome, and placed onto glass slides. Between samples, the microtome blade was replaced to prevent cross contamination. One 3μm section was stained with haematoxylin and eosin and examined histologically to determine areas of tumour with minimal stromal and inflammatory cell contamination. The selected area was removed from the 10μm section using a sterile scalpel to scrape the tissue into a microfuge tube. For each NSCLC case to be analysed 10μm sections from lung tissue or lymph node, previously examined histologically and shown not to contain tumour ('normal'), were also cut and placed directly into microfuge tubes for DNA extraction.

3.3.2 DNA extraction from paraffin embedded tissue

DNA was extracted from the paraffin embedded tissue sections using the method of Levi and co-workers (1991), with slight modifications. For sections isolated from slides the method of DNA extraction was as follows; samples were microfuged briefly to pellet the tissue section and 400μl of lysis buffer and 25μl of proteinase K (20 mg/ml) added (see appendix A). The samples were incubated for a period of 1 hour at 55°C, gently vortexed then incubated for a further 3 hours. Following incubation, sterile EDTA, pH 8.0 was added to a final concentration of 1 mM to prevent bacterial growth in the sample and prevent digestion of DNA by nucleases.

Sections of normal tissue were dewaxed before following the above extraction protocol.

To dewax, 1 ml of xylene was added, and the section vortexed and left to stand for 2 minutes at room temperature (RT). The samples were centrifuged at 16,000 rpm for 2 minutes, the supernatant discarded and the pellet resuspended in 1 ml of xylene and centrifuged as previously described. Finally, 1 ml of absolute ethanol was added and samples left for 2 minutes at RT. The samples were centrifuged again, the supernatant removed, and this step was repeated. The pelleted tissues were air-dried and the protocol for DNA extraction outlined above was followed. All samples were aliquoted and stored at –20°C.

3.3.3 Determination of DNA concentration using ethidium bromide plates

Ethidium bromide plates were prepared by adding 1g of agarose into 100 ml of 1x TBE buffer. Agarose was dissolved in the TBE by slowly microwaving. Once the solution had cooled to approximately 50°C, ethidium bromide solution (an intercalating dye that fluoresces under ultraviolet light) was added to give a final concentration of 0.1 mg/ml. The solution was poured into petri dishes and allowed to cool until set. Plates were stored at 4°C until required.

To determine the concentration of a DNA sample, 1µl of the sample and 1µl of each concentration of a standard DNA, ranging from 3 ng/µl to 125 ng/µl, were pipetted onto the plate. Plates were incubated at room temperature for 25 minutes and viewed under ultraviolet light. Concentrations of unknown samples were determined by eye by comparing with florescence of the known standards.

3.3.4 Analysis of polymorphic sites within *APC* and *MCC* genes for loss of heterozygosity

Loss of heterozygosity (LOH) in the *APC* and *MCC* genes was determined using PCR amplification of intragenic sequences containing known polymorphisms. Following PCR amplification restriction fragment length polymorphisms (RFLPs) in the 3' untranslated region (UTR) (Heighway *et al.*, 1991) and exon 11 of the *APC* gene (Groden *et al.*, 1991) and the 3' UTR of the *MCC* gene (Heighway *et al.*, 1992) were analysed using

restriction endonucleases *Ssp1*, *Rsa1* and *Mae111*, respectively. All enzymes were purchased from Boehringer Manheim. A fourth polymorphism, a length polymorphism in exon 10 of the *MCC* gene (Miyoshi *et al.*, 1992a) was also analysed.

3.3.4.1 Amplification of intragenic sequences containing known polymorphisms

Intragenic sequences were amplified using polymerase chain reaction (PCR). PCR reactions were carried out in 100µl volumes. Each reaction mix consisted of 0.1-1µg DNA, 2.5 units *Taq* polymerase (Gibco BRL), 0.2 mM of each dNTP (dATP, dCTP, dTTP, dGTP) (Pharmacia Biotech.), 1.5 mM $MgCl_2$ (Gibco BRL), 10µl of 10 X PCR buffer (200 mM Tris-HCl pH 8.0, 500 mM KCl. Gibco BRL) and 0.5µM of each primer (Oswell DNA Services). PCR was performed in a Hybaid Omni Gene DNA thermocycler. Primers used for the amplification of the intragenic sequences have been published (Heighway *et al.*, 1991,1992, Groden *et al.*, 1991 and Miyoshi *et al.*, 1992a) with the exception of a primer used for the amplification of the 3' UTR of the *APC* gene. Here the previously published upstream primer was replaced with a new primer to produce a smaller PCR product that was easier to amplify from archival paraffin-embedded tissue. Primers, specific PCR annealing conditions and cycle number for each set of primers is shown in table 4. The following standard PCR cycling program was used in all cases; initial denaturation step of 5 minutes at 94°C followed by a cycling program of denaturation at 94°C for 1 minute, annealing conditions see table 4, and an extension step of 72°C for 1 minute with the exception of *MCC* exon 10 which had a 2 minutes extension time. A final extension step of 72°C for 9 minutes was also included.

Ten microlitres (approximately 0.2 µg) of each PCR product was electrophoresed through a 3% agarose gel (see below) to ensure amplification of the template.

3.3.4.2 Agarose gel electrophoresis

To prepare a 3% agarose gel, 4.5 grams of agarose was weighed out into a flask and 150 ml of 1xTBE buffer added. The agarose solution was gently heated in a 600W microwave for 2 minutes or until all the agarose powder was dissolved. A magnetic bar was added to

the agarose solution and the flask was placed on a magnetic stirrer for 5 minutes. Ethidium bromide was added to the agarose to give a final concentration of 0.1 mg/ml. The agarose gel was poured into a cast and left to set at 4°C.

10µl of each PCR product was mixed with 1µl of 10X gel loading buffer (Appendix A) and the solution loaded into the wells of the agarose gel. Gels were run at 100 V in 1X TBE buffer. Following electrophoresis PCR product was visualised by illuminating the gel with ultraviolet light.

Table 4: Amplification of restriction fragment polymorphisms in *APC* and *MCC* genes.

POLYMORPHISM	PRIMERS 5'-3'	PCR PARAMETERS AnnealingTemp./Time	Program cycles
APC 3' UTR	U – ATGACCACCAGGTAGGTGTATT D – GAAGAGACTGCAATGTCTAAGAA* OR D –GCATTAAGAGTAAAATTCCTCTTAC	55°C 30 seconds. OR *54°C 30 seconds.	30 30
APC EXON 11	U-GGACTACAGGCCATTGCAGAA D –GGCTACATCTCCAAAAGTCAA	59°C 1 minute.	34
MCC 3' UTR	U-CCAATGAAACTTCGCTTTAATCAG D –CTGGATACAGTCCACAATGACAC	52°C 1 minute.	36
MCC EXON 10	U-CACTTCTACCCTGAAGTAGCTCC D –ATGACCTCCTGACCATAACCTTG	55°C 1 minute.	30

*New Primer, U-upstream primer, D-downstream primer

3.3.4.3 Restriction endonuclease cleavage of PCR products

Approximately 0.2μg of each PCR product was incubated in a total volume of 20μl containing 2 units of the appropriate restriction endonuclease and 1 x restriction enzyme buffer for 2 hours at the recommended temperature. In the case of *APC* exon 11 and *MCC* 3' UTR, a control DNA fragment (internal control) containing the appropriate restriction site was included to ensure restriction enzyme cleavage. In these cases the incomplete cleavage of the PCR product may lead to incorrect interpretation of results as the sample would be scored as either non-informative or as informative with loss of the second allele.

3.3.4.4 Analysis of polymorphic sites

Following restriction enzyme cleavage, the tumour and 'normal' PCR product pairs were electrophoresed through a 3% agarose gel in 1x TBE running buffer until the different size alleles were separated. Molecular weight markers (marker V, Boehringer Manheim) were included on each gel. Samples were scored as informative (polymorphic between alleles) or non-informative (no-polymorphism between alleles). LOH in informative cases was defined by a reduction in the intensity, as determined by eye, of an allele band of the tumour in comparison with its corresponding normal. PCR product and RFLP sizes are shown in table 5.

Table 5: Analysis of restriction fragment polymorphisms.

Polymorphism	Restriction enzyme	PCR Product Size (bp)	Allele size (bp) Following enzyme cleavage	Internal Control (bp)
APC 3' UTR	*Ssp1*	850	580 +270 A1 580 +135 +135 A2	
APC 3' UTR	*Ssp1*	318	270+48 A1 135+135+48 A2	
APC exon 11	*Rsa1*	132	132 A1 87 + 45 A2	PUC 19 PCR product size 103 Cleavage 71 32
MCC 3' UTR	*Mae111*	210	210 A1 189 +21 A2	*APC* 3' UTR PCR product size 850 Cleavage 700 150
MCC exon 10	Not applicable	175 161	175 A1 161 A2	

IC- Internal Control

A1/A2- Allele bands 1 or 2.

3.3.5 Single strand conformational polymorphism analysis of the MCR of the *APC* gene

All cases showing LOH at 5q21 were screened for mutations in the mutation cluster region (MCR) of the remaining allele of *APC*. The method used was single strand conformational polymorphism (SSCP); a method designed to detect unknown point mutations. The method was originally described by *Orita et al.*, (1989) and involves amplification of DNA using PCR and electrophoresis of denatured product on a polyacrylamide gel under non-denaturing conditions. The electrophoretic mobility of the single stranded DNA is dependent on the size and the folded secondary structure of the DNA. The secondary structure of the DNA is stabilised by intra-strand nucleotide interactions determined by the base sequence of the strand.

3.3.5.1 Amplification of DNA fragments spanning the mutation cluster region

The MCR of the *APC* gene lies between codons 1268 and 1513 (nucleotides 3717-4614) (Miyoshi *et al.*, 1992). Amplification of DNA fragments spanning the MCR was carried out using polymerase chain reaction (PCR). Primers used to span the MCR were as previously reported (Curtis *et al.*, 1994) but modified to produce PCR products less than 300 bp long (suitable for amplification of DNA extracted from paraffin embedded tissue). Nucleotide location, PCR product size, primers and PCR annealing temperatures are noted in table 6. Cycling conditions consisted of an initial denaturation step of 5 minutes at 94°C, followed by 36 cycles of 30 seconds denaturation at 94°C, 30 seconds at appropriate annealing temperature, and 60 seconds extension at 72°C, with a final extension step of 10 minutes at 72°C.

To ensure that the PCR reaction had been successful, PCR products were analysed on 3% agarose gels using the method described in section 3.3.3.2.

Table 6: Sections of *APC* mutation cluster region for SSCP analysis.

APC MCR SECTION	NUCLEOTIDE LOCATION	PRODUCT SIZE (bp)	PRIMERS 5'-3'	Annealing Temperature °C
1	3735-3991	256	U-TGGAACTTCGTCACAGGAT D-AAGTGGTCAGCCTCAAAAGG	58
2	3888-4181	292	U-TCAGACGACACAGGAAGCAG D-GTACATCTGCTAAACATGAGTGGG	56
3	4150-4441	291	U-CAGGAGACCCCACTCATGTT D-CAGCATTTACTGCAGCTTGC	56
4	4408-4612	203	U-AGAGTGGACCTAAGCAAGCT D-ATTTTCCTGAACTGGAGGC	55

U-Upstream Primer, D-Downstream Primer

3.3.5.2 Single strand conformational polymorphism gels

PCR products were denatured by the addition of 1µl of denaturing solution (Appendix A) to 5µl (approximately 0.1µg) of PCR product. Samples were incubated for 5 minutes at 50°C, then 3µl of stop solution (Appendix A) was added to each sample. Nine microlitres of sample was loaded onto a non-denaturing 0.7% MDE gel (AT Biochem, PA. USA) containing 5% glycerol and 1xTBE. The gel was run for 2-3 hours in 1x TBE running buffer at 25°C at 25W using a LKB Bromma 2001 vertical electrophoresis unit. Tumour and normal DNA pairs were loaded side by side, and DNA from colorectal adenocarcinomas previously shown to harbour mutations in the MCR of *APC* were included as positive control on some gels.

3.3.5.3 Silver staining and gel analysis

Gels were silver stained according to the manufacturer's instructions (Bio-Rad Laboratories Ltd., UK). This involved fixing the gels by washing in 40% methanol / 10% glacial acetic acid (v/v) for 1 hour, followed by two 30 minute washes in 10% ethanol / 5% glacial acetic acid (v/v). The gels were then incubated at RT on a rocking platform for 10 minutes in oxidising solution, which contained potassium dichromate and nitric acid, and washed in double distilled water (DDW) for 2 minutes. The DDW water was removed and the washing repeated at least 12 times until the yellow coloration had been removed from the gels. Gels were incubated at RT on a rocking platform in silver nitrate solution for 30 minutes. The gels were then washed for 2 minutes in DDW and transferred to developer solution, which contained sodium carbonate and paraformaldehyde. Initially 200 ml was added and immediately poured off. A further 200 ml was added and the gels incubated at RT until brown DNA bands appeared on the gel, usually after 5 minutes. Developer was then removed and 5% glacial acetic acid (v/v) added for approximately 15 minutes to fix the gel. Gels were then dried on 3M Whatman paper and laminated. Gels were examined for the presence of band mobility shifts between normal and tumour DNA samples. Mobility shifts were seen in all positive control samples.

3.3.6 APC immunohistochemistry

All cases showing LOH at 5q21 were analysed by immunohistochemistry using a rabbit polyclonal antibody raised to the carboxyl terminus (C-terminal) of the APC protein to determine whether full length APC protein was being translated from the remaining *APC* allele. This APC antibody was produced within the department and is fully described in Midgley *et al.*, 1997. Three micron sections from paraffin-embedded tissues were dewaxed and hydrated by incubating for 10 minutes in xylene followed by 2-minutes each in absolute alcohol, and graded aqueous alcohol (95% and 70%). Slides were washed for 2 minutes in tap water and incubated for 15 minutes in 3% (v/v) hydrogen peroxide solution diluted in water to block endogenous peroxidase activity. Slides were washed in running tap water for 5 minutes and incubated in 2 changes of phosphate buffered saline (PBS) with 0.1% (v/v) Tween-20 for 5 minutes. Sections were incubated for 20 minutes with 100µl of PBS containing 20% normal donkey serum (Dako, UK) and 0.1% Tween-20 (NDSPT). Following incubation slides were washed with PBS/ 0.1% Tween-20 and then 100µl of APC C-terminal antibody, diluted 1/50 (a dilution previously determined) in NDSPT, was added to each tissue section. Slides were incubated overnight at 4°C in a humidifying chamber. Slides were then washed for three 5-minute washes in PBS/ 0.1% Tween-20 and sections were incubated for 30 minutes in a humidifying chamber with 100µl of horseradish peroxidase conjugated donkey anti-rabbit secondary antibody (Dako, UK), diluted 1/400 in PBS/0.1% Tween-20. Again slides were washed for three 5 minute washes PBS / 0.1% Tween-20. For detection of immunoperoxidase activity, freshly prepared diaminobenzidine (DAB) solution (Appendix A) was applied to the slides for 5 minutes. Slides were washed in water for a further 5 minutes and nuclei were counter-stained with Mayer's haematoxylin for 5 seconds. After a further wash in running tap water, the sections were transferred through graded aqueous alcohol solutions to absolute ethanol and then through 3 changes of xylene. Cover slips were mounted using DPX (BDH, UK) mounting fluid. Control slides from 'normal' lung tissue were included

as 1) a positive control slide to confirm APC staining 2) a negative control which was incubated with NDSPT instead of APC antibody.

3.4 Results

3.4.1 LOH at 5q21

Ninety-nine NSCLC cases were examined for LOH at polymorphic sites within the *APC* and *MCC* genes. NSCLC cases consisted of 33 consecutive SCC, 59 consecutive ADC and 7 additional non-randomly selected ADC of bronchial origin. The latter cases were included to enhance cases for statistical analysis between ADC of parenchymal and bronchial origin (see section 1.2.1.2 and 3.3.1) but were not included in the initial analysis of frequency of LOH in our consecutive cohort, as seen in table 7a.

Four genetic polymorphisms in the *APC* and *MCC* genes were analysed. Following PCR amplification and subsequent incubation with the appropriate restriction endonuclease (if required), the polymorphic sites within tumour and normal tissue pairs for each NSCLC cases were viewed under ultraviolet light following electrophoresis through a 3% agarose gel containing ethidium bromide (see section 3.3.4).

Normal tissue samples from each NSCLC cases were viewed to determine whether polymorphic sites were non-informative, i.e. no polymorphic variation between alleles, or informative i.e. showed polymorphic variation between alleles. Tumour and normal tissue pairs of cases shown to be informative were then scored as informative retained, the tumour sample retained both alleles, or informative with LOH, one allele was absent or was noted to be of decreased intensity when compared to the corresponding allele within the normal tissue sample. Example of cases showing non-informative, informative retained and informative with LOH for each site are shown in figures 9 and 10.

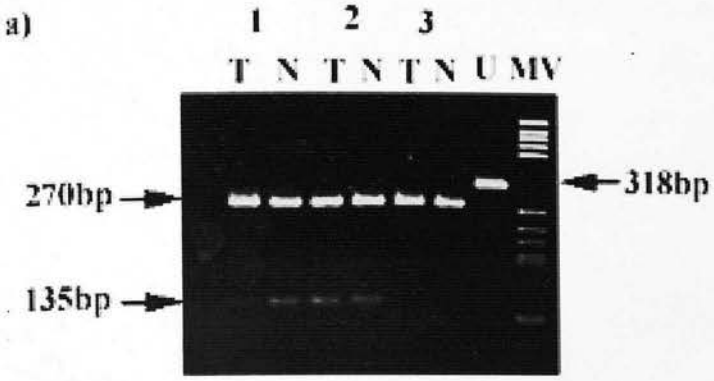
Unlike in the analysis of tumour cell lines, DNA extractions from tumour tissue contain non-tumour DNA derived from stromal tissue and infiltrating inflammatory cells. This can make the interpretation of results difficult and subjective. To try and eliminate misinterpretation, Dr VJ Bubb and myself independently scored results. In cases where there was disagreement, DNA was re-extracted from the tissue trying to further minimise the amount of stromal material. As a further precaution, only cases informative at two or more loci were subject to statistical analysis.

Figure 9: Analysis of restriction fragment length polymorphisms (RFLP) within the *APC* gene

a) This figure shows the analysis of a polymorphic site within the 3` untranslated region of the *APC* gene. DNA extracted from tumour (T) and normal (N) tissue for each NSCLC case was subject to PCR to amplify a 318 bp product which encorporated the polymorphic site. PCR products were incubated with the restriction endonuclease *Ssp1* to identify polymorphic variants which would yield fragments of 270 + 48 bp or 135 + 135 + 48 bp. Following incubation with *Ssp1*, PCR products were electrophoresed on a 3% agarose gel. PCR product not incubated with *Ssp1* was also loaded on the gel to indicate the migration of undigested PCR product, this lane is indicated as "U". The 48 bp restriction fragment can not be visualised on this gel.
Case 1 is an individual who is heterozygous for this polymorphism and exhibits loss of one *APC* allele in the tumour sample (loss of heterozygosity). Case 2 is an individual who is heterozygous for this polymorphism and both alleles are retained in the tumour and normal samples (informative retained).Case 3 is an individual who is homozygous (non-informative). MV indicates molecular weight marker V (Boehringer Manheim).

b) This figure shows the analysis of a polymorphic site within exon 11 of the *APC* gene. PCR amplification of DNA samples yields a 132 bp product which when incubated with the restriction endonuclease *Rsa1* identifies the two polymorphic variants yielding fragments of 132 bp or 87 + 45 bp. Internal controls to ensure enzymatic cleavage were required (see section 3.3.4.3). PCR amplification of the internal control DNA yields a product of 103 bp that is cleaved by *Rsa1* to give fragments of 71 + 32 bp. The 32 bp fragment can not be visualised on this 3% agarose gel.
Following incubation with *Rsa1*, PCR products from tumour (T)and normal (N) tissue for each case were electrophoresed on a 3% agarose gel. PCR product not incubated with *Rsa1* were loaded onto the gel to indicate the migration of the 132bp PCR product, indicated on the figure as U. Lanes C1 and C2 correspond to PCR amplification of the internal control uncleaved and cleaved with *Rsa1* respectively.
Case 1 is an individual who is heterozygous for the polymorphism and exhibits loss of one of the *APC* alleles within the tumour sample. Cases 2 and 3 are individuals who are informative for the polymorphism and both alleles are retained in the normal and tumour samples. MV indicates molecular weight marker V.

Figure 9: Analysis of polymorphic sites within the *APC* gene



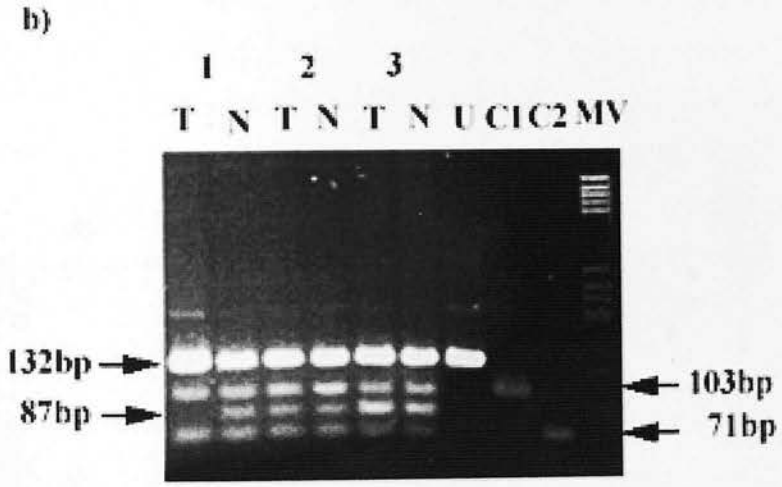


Figure 10: Analysis of polymorphic sites within the *MCC* gene

a) This figure shows the analysis of a RFLP within the 3' untranslated region of the *MCC* gene. DNA extracted from tumour (T) and normal (N) tissue for each NSCLC case was subject to PCR to amplify a 210 bp product, which encompassed the polymorphic site. PCR products were incubated with the restriction endonuclease *Mae111*, which identifies the two polymorphic variants yielding fragments of 210 bp or 189 +21 bp. Internal controls to ensure enzymatic cleavage were required (see section 3.3.4.3). Amplification of the internal control yielded a PCR product of 850bp, which when incubated with MaeIII yielded fragments of 700 and 150 bp. Lane marked C in the figure conatins both the PCR product from the amplification of the RFLP in the MCC gene and the internal control. Following incubation with *Mae111*, PCR products were electrophoresed through a 3% agarose gel. The 21 bp fragment can not be visualised on this gel.

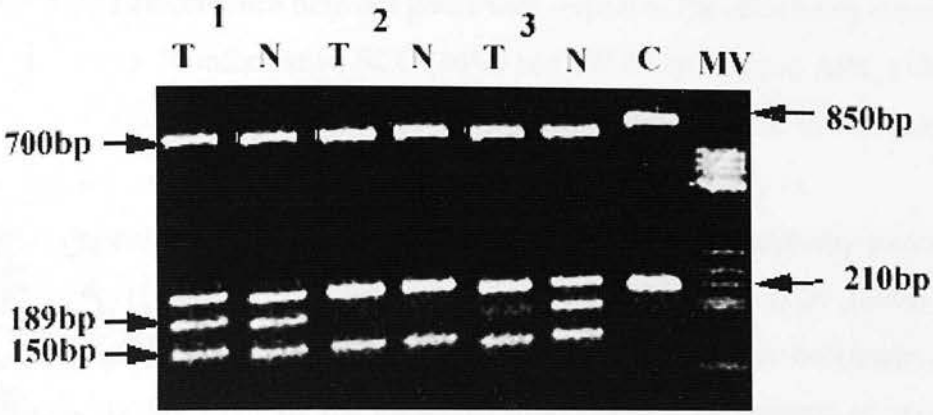
   Case 1 is an individual who is heterozygous for the polymorphism and both alleles are retained in the tumour and normal tissue. Case 2 is an individual who is homozygous for the polymorphism (non-informative). Case 3 is an individual who is heterozygous for this polymorphism and exhibits loss of one of the *APC* alleles within the tumour sample (loss of heterozygosity). MV indicates molecular weight marker V (Boehringer Manheim).

b) This figure shows the analysis of a length polymorphism within exon 10 of the *MCC* gene. Amplification of a fragment of DNA encompassing the polymorphic site identifies the two polymorphic variants, yielding products of 175 bp or 161 bp.

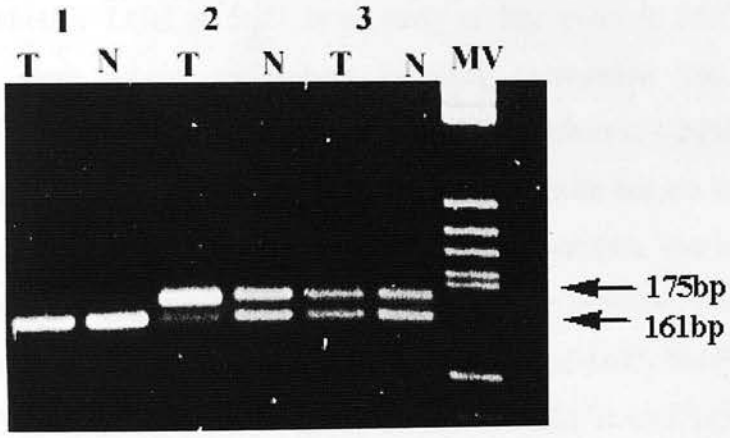
   Case 1 is an individual who is homozygous for this polymorphism. Case 2 is an individual who is heterozygous for the polymorphism and exhibits loss of one of the *APC* alleles within tumour tissue. Case 3 is an individual who is heterozygous for the polymorphism and both *APC* alleles are retained in tumour and normal tissue. MV indicates molecular weight marker V (Boehringer Manheim).

Figure 10: Analysis of polymorphic sites within the *MCC* gene

a)



b)



All case data and histological information is shown in Appendix C. Summary of results is shown in table 7 a-d.

Sixty-eight of the 92 consecutive NSCLC cases were informative at two or more loci. Cases that contained informative polymorphisms in both the *APC* and *MCC* genes exhibited no discordance between genes with respect to the allele being retained or lost.

Eleven out of 24 informative SCC (46%) and 16/ 44 informative ADC (36%) exhibited LOH. Statistical analysis indicated no significant difference in the frequency of LOH between SCC and ADC (χ^2=0.0.582, df=1; P=0.446). See table 7a.

Within the consecutive ADC series with the additional non-randomly selected bronchial ADC, 5/9 (56%) of informative bronchial ADC whilst 6/20 (30%) informative parenchymal ADC exhibited LOH. Nineteen ADC could not be clearly described as arising in the bronchus or the parenchyma. Of these cases, termed of uncertain origin, 8/19 (42%) showed LOH. Fisher's exact test was applied to these results and demonstrated that there was no statistically significant difference between ADC of parenchymal or bronchial origin with respect to LOH at 5q21 (Two-tailed analysis; p=0.237). See table 7b.

To determine whether LOH at 5q21 is an early or late event in NSCLC, the cases analysed above were subject to analysis following segregation into tumour stage according to the international staging system, as outlined in section 1.2 (Mountain 1997). No significant difference was seen between tumour stages with respect to LOH at 5q21 (Fisher's exact test; Two-tailed test; p>0.2 in all cases). Therefore, one may conclude at this point, that loss of 5q21 is not associated with tumour progression. When the NSCLC group was subdivided into the histological groups ADC and SCC, results revealed that the frequency of LOH within the SCC group remained similar at each tumour stage, and statistical analysis showed that there was no significant difference between tumour stage and LOH, the same was shown for the ADC group, Fisher's exact test, Two-tailed test; p>0.2. See table 7c. No significant difference was seen in the frequency of LOH between SCC and ADC groups at each stage, Fisher's exact test p>0.2.

Following the subdivision of ADC into site of origin, results revealed that LOH was not associated with increasing tumour stage in the parenchymal or bronchial subgroups, as determined by Fisher's exact test (two-tailed test; $p>0.2$ in all cases). The group of tumours designated of uncertain origin followed the same pattern of loss as the ADC of parenchymal origin. See table 7d.

Table 7 a) Analysis of consecutive NSCLC cases for LOH

Histology	No. of informative cases	No. of informative cases exhibiting LOH (%)
SCC	24	11 (46%)
ADC	44	16 (36%)

Ninety-nine consecutive NSCLC cases were analysed for LOH, 68 cases were shown to exhibit loss of heterozygosisty at more than one loci. Statistical analysis indicated no significant difference in the frequency of LOH between SCC and ADC (χ^2=0.582, df=1; p=0.446).

Table 7 b) Analysis of ADC cases for LOH following subclassification into site of origin

Subclassification of ADC	No. of informative cases	No. of informative cases exhibiting LOH (%)
Parenchymal	20	6 (30%)
Bronchial	9	5 (56%)
Uncertain origin	19	8 (42%)

Fourty-four ADC cases identified in the initial consecutive NSCLC cohort, plus 7 non-consecutive bronchial ADC, which were selected to improve statistical analysis, were analysed. Four of the seven bronchial ADC were informative, of which three exhibited LOH. Statistical analysis indicated no significant difference in the frequency of LOH between ADC of parenchymal and bronchial origin, Fisher's exact test, two-tailed test; p=0.237.

Table 7 c) Analysis of tumour stage and LOH in ADC and SCC subtypes

Tumour stage	Total no. of cases exhibiting LOH / informative cases	No. of SCC exhibiting LOH / informative cases	No. of ADC exhibiting LOH / informative cases
Ia +Ib	16/37 (43%)	5/9 (56%)	11/28 (39%)
IIa + IIb	8/23 (35%)	3/10 (30%)	5/13 (38%)
IIIa + IIIb	6/12 (50%)	3/5 (60%)	3/7 (43%)

All cases were analysed to determine if there was any significance difference between the frequency of LOH with tumour stage for the ADC and SCC groups. No significant difference in LOH between tumour stages was seen for SCC, nor for ADC, Fisher's exact test, Two-tailed test; $p>0.2$ in all cases.

Table 7 d) Analysis of tumour stage and LOH following subdivision of ADC into parenchymal and bronchial subtypes.

Tumour stage	No. of ADC of uncertain origin exhibiting LOH/ informative cases	No. of ADC of bronchial origin exhibiting LOH/ informative cases	No of ADC of parenchymal origin exhibiting LOH/ informative cases
Ia +Ib	2/10 (20%)	5/8 (62%)	4/11 (36%)
IIa + IIb	3/7 (43%)	0	2/4 (50%)
IIIa + IIIb	1/3 (33%)	0/1 (0%)	2/4 (50%)

No statistically significant difference was seen in frequency of LOH between tumor stages, in either the bronchial or parenchymal ADC groups, was seen as determined by Fisher's exact test, two-tailed test; $p>0.2$ in all cases. Comparison of LOH revealed no significant difference in the frequency of LOH between ADC arising in the bronchus or parenchyma at each tumour stage, Fisher's exact test, two-tailed test; $p= >0.2$ in all cases.

3.4.2 Mutational analysis of the *APC* gene

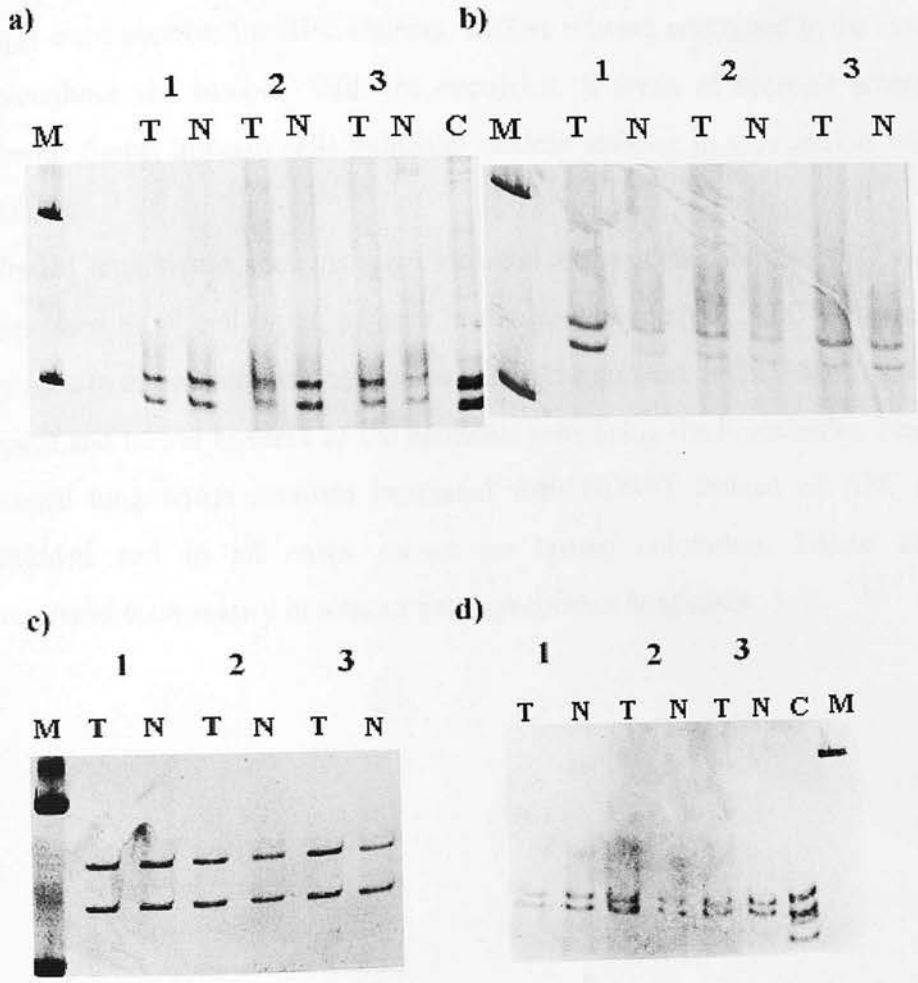
NSCLC cases shown to exhibit LOH at 5q21 (n=30) were subject to mutation analysis. Single strand conformation polymorphism analysis was used to screen for mutations within the MCR of the *APC* gene. PCR was used to amplify the MCR, in several overlapping fragments, using DNA extracted from tumour and normal tissue of each NSCLC case. For each MCR fragment, PCR products from tumour and normal tissue were loaded side by side onto an SSCP gel, subject to electrophoresis and then silver stained to visualise the PCR products (see section 3.3.5).

As positive controls, DNA extracted from colorectal ADC cases known to harbour mutations within the MCR of the *APC* gene, were also amplified by PCR and included on the gel. In all cases, positive controls showed a shift in band mobility when compared to the NSCLC test samples. No shift in band mobility was seen between tumour and corresponding normal tissue from individual NSCLC cases, or between all NSCLC cases. Examples of SSCP gels are shown in figure 11.

Figure 11: Single strand conformational polymorphism analysis of the mutation cluster region of *APC*

Primers were designed to amplify the mutation cluster region (MCR) of the *APC* gene. Following amplification of a fragment of the MCR by PCR, PCR products were electrophoresed though an SSCP gel and subsequently silver stained. For each NSCLC case, previously shown to exhibit LOH at 5q21, tumour (T) and normal (N) tissue was compared. Figures 11a –11d are examples of the analysis of the MCR fragments 1-4 respectively. On each gel, 3 NSCLC cases are represented (1-3). Control samples, shown in the figure in lanes marked C, were also analysed, these cases were tissue from colorectal cancer cases previously shown to harbour a mutation within that region of the MCR. A shift in band mobility is noted for both of these positive control cases. No shifts in band mobility are noted in any of the tumour or normal tissue pairs or between NSCLC cases. Lane M denotes a molecular weight marker V (Boehringer Manheim).

Figure 11: Single strand conformational polymorphism analysis of the mutation cluster region of *APC*



3.4.3 APC immunohistochemistry

Tumour tissue from NSCLC cases shown to exhibit LOH at 5q21 (n=30) were subjected to immunohistochemistry using a polyclonal antibody raised to the carboxyl terminal of the *APC* gene (Midgley *et al.*, 1997). NSCLC were scored as positive by the presence of a brown coloration within the cell, reflecting the diaminobenzidine substrate, which was used for visualisation of the immunohistochemical reaction. All 30 cases showing LOH at 5q21 were positive for APC staining. Diffuse staining was noted in the cytoplasm of cells throughout the tumour, with the exception of areas of necrosis where staining was absent. Some tumour cells exhibited nuclear staining in conjunction with cytoplasmic staining.

Normal lung tissue sections were included as positive controls. *APC* was ubiquitously expressed in all cell types, as seen in the tumour samples. APC was present within the cytoplasm of cells and in some cases within the nucleus. Intense staining was noted at the apical and lateral borders of the epithelial cells lining the bronchioles. Negative controls, normal lung tissue sections incubated with NDSPT instead of APC antibody, were included and in all cases shown no brown coloration. Figure 12 shows APC immunohistochemistry in tumour and non-tumour lung cases.

Figure 12: APC protein expression in normal and tumour tissue of the lung

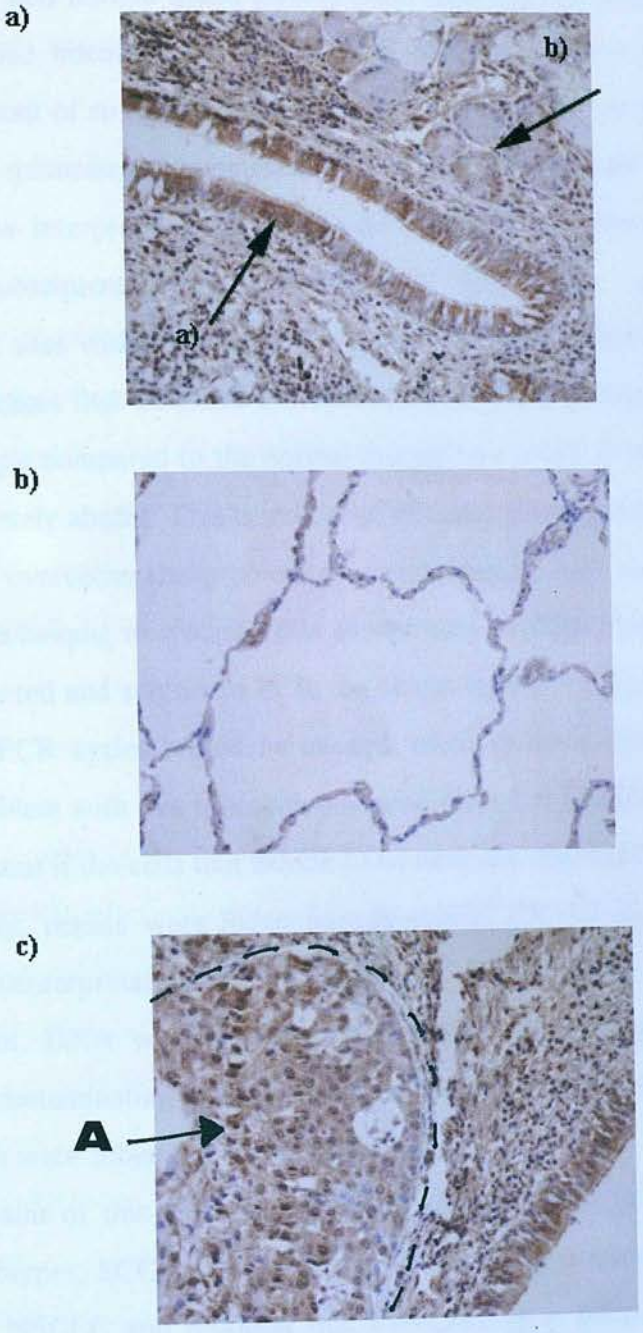
Immunohistochemical staining of formalin fixed paraffin embedded lung tissue with a polyclonal antibody raised to the carboxyl terminal of the APC protein. The method used was indirect immunoperoxidase using diaminobenzidine substrate for visualisation. Slides were lightly counter stained with haematoxylin. APC protein was detected in all cells types; diffuse cytoplasmic staining was noted with occasional nuclear staining within some cells.

a) This section shows APC protein expression in the bronchial epithelium. Arrow a indicates intense apical and lateral staining of the epithelial cells lining the bronchioles (cuboidal ciliated cells and columnar Clara cells). Arrow b indicates staining within the submucosal gland. Magnification x 200.

b) APC protein expression in alveolar epithelial cells (type I and type II Clara cells). Magnification x 400.

c) APC protein expression in a bronchial adenocarcinoma (area a). This case showed loss of heterozygosity (loss of one *APC* allele) as determined by restriction fragment polymorphism analysis. Staining of the tumour indicates expression of full-length protein from the remaining *APC* allele. Magnification x 400.

Figure 12: APC protein expression in normal and tumour tissue of the lung



3.5 Discussion

Polymorphic sites within the *APC* and *MCC* genes were analysed to determine the frequency of LOH at chromosome 5q21 in a series of NSCLC cases. DNA was extracted from tumour and normal tissue derived from each NSCLC case. Areas of tumour tissue were examined histologically and an area of tumour tissue carefully selected with a minimal amount of stromal tissue and infiltrating inflammatory cells. Tissue selection was necessary to minimise contamination of the tumour DNA with non-tumour DNA as this can make the interpretation of results difficult and subjective. Following PCR, and if required, subsequent treatment with the appropriate restriction endonucleases, polymorphic sites within the tumour and normal samples were analysed for LOH. In the majority of cases that exhibited LOH a reduction in the intensity of an allele band in the tumour sample compared to the normal sample was noted. It was rare for the allele band to be completely absent. This is indicative of contaminating non-tumour tissue within the sample. To overcome the problem of contamination, one could use a more stringent dissection technique; microdissection can be used to select single cells from which DNA can be extracted and subject to PCR. To obtain enough PCR product for analysis, a high number of PCR cycles would be needed, often in these circumstances nested PCR is used. A problem with this protocol may arise if there is heterogeneity within the tumour. This may occur if the cells that exhibit LOH have not reached clonal dominance.

In this study, results were independently scored (Dr VJ Bubb and myself) to try to eliminate misinterpretation due to non-tumour contamination. In cases where there was disagreement, DNA was re-extracted from the tissue trying to further minimise the amount of contaminating material. As a further precaution, only cases informative at two or more loci were subject to statistical analysis.

The initial aim of this study was to determine the frequency of LOH at 5q21 in the NSCLC subtypes; SCC and ADC. Previous studies have analysed the frequency of LOH at 5q21 in NSCLC and reported that LOH occurs in between 20-70% of cases. The variation in frequency of loss seen between these studies suggested that LOH may be more common in different NSCLC subtypes or associated with different tumour stages,

as these were variables between these studies.

This study has shown that in this consecutive cohort of NSCLC cases, LOH at 5q21 is a frequent event. In the SCC group LOH was detected in 11/24 (46%) informative SCC cases and 16/44 (36%) informative ADC cases (Table 7a). Statistical analysis reveals that there is no significant difference in the frequency of loss in ADC and SCC subtypes. This result is supported by two smaller studies (Ashton-Rickardt *et al.*, 1991; Tsuchiya *et al.*, 1992). Conversely, two studies have reported that LOH at 5q21 occurs more frequently in SCC than ADC, in both studies results were statistically significant (Fong *et al.*, 1995; Sanz-Ortega *et al.*, 1999).

The differences seen between cohorts may have been a reflection of the number of early and late stage tumours. However, in this study, SCC of all tumour stages exhibited similar frequencies of LOH (stage I: 5/9 (56%); stage II: 3/10 (30%); stage III: 3/5 (60%)), as seen in table 7c. Stataistical analysis of the data revealed that there was no association between tumour stage and frequency of LOH, Fisher's exact test. Two-tailed test; $p>0.2$ in all cases, and suggested that LOH at 5q21 is not associated with tumour progression. This result is in keeping with two recently published studies (Sanz-Ortega *et al.*, 1999; Wistuba *et al.*, 1999) but is contrary to studies by Fong and Co-workers (1995a).

Analysis of LOH and tumour stage for ADC cases also revealed that the frequency of LOH did not increase with tumour progression, stage I: 11/28 (39%), stage II: 5/13 (38%) and stage III: 3/7 (43%), data shown in table 7c..Using Fisher's exact test, the frequency of LOH at 5q21 between tumour stages was shown not to be statistically significant, two-tailed test; $p>0.2$ in all cases, thus indicating that LOH at 5q21 is not associated with tumour progression.

Subdivision of ADC into site of origin (see table 7b) revealed no significant difference in the frequency of LOH occurring in ADC of parenchymal or bronchial origin (Fisher's exact test, two-tailed test; $p=0.237$). Statistical analysis was then used to determine whether the frequency of LOH in parencymal and bronchial subgroups varied with tumour stage (data shown in table 7d), no significant difference was determined, Fisher's

exact test, two-tailed test; p >0.2 in all cases. Analysis of the frequency of LOH exhibited by bronchial and parenchymal ADC at each tumour stage revealled that the frequency of loss was similar (see table 7d) and statistical analysis revealed no significant difference, Fisher's exact test p>0.2 in all cases. Therefore, through the analysis of individual ADC subgroups, it appears that LOH at 5q21 is not associated with tumour progression within either subgroup and that the frequncy of LOH exhibited by the two subgroups is not significantly different.

In summary, this study has shown that there is no statistically significant difference in the frequency of LOH at 5q21 in the NSCLC histological subgroups ADC and SCC and that LOH is not associated with tumour progression in these groups. These results suggest that with respect to LOH at 5q21, ADC and SCC may follow the same genetic pathway to tumourigenesis and that LOH of heterozygosity at 5q21 does not promote tumour progression. Therefore, one may speculate that LOH at 5q21 may be important in the initiation of tumourigenesis. Analysis of ADC cases of bronchial and parenchymal origin revealed that the frequency of LOH is not significantly different between these subgroups and that frequency of loss does not increase significantly with tumour stage. This suggests, as with ADC and SCC, that LOH is not associated with tumour progression in either of these subgroups and that LOH at 5q21 may be important in the initiation of tumourigenesis.

There are currently no publications, which identify the point at which LOH at 5q21 occurs. It is also still unclear as to why the frequency of LOH at 5q21 varies to such an extent. Despite carrying out statistical analysis using the Fisher's exact test which allows for small numbers, I think that it may be worth increasing the numbers of cases within this study, in particular the number of later stage tumours. This will be difficult as in most cases NSCLC are detected and surgical procedures are immediately implemented before the tumours can progress. Alternatively, a study of advanced tumours may be carried out through the production of xenografts using murine hosts.

In NSCLC, the region of chromosomal loss at 5q21 spans 3-5 megabases and harbours numerous genes including the tumour suppressor gene *APC* and the *MCC* gene (see

section 1.3.3.3a). These two genes lie approximately within 500 kb of each other. Initially, *MCC* was thought to also be a tumour suppressor gene as mutations were detected in a series of colorectal cancer cases (Kinzler *et al*, 1991a). Subsequent studies where undertaken to determine the frequency of gene mutations in a larger series of colorectal cancer cases and these studies revealed no mutations in the *MCC* gene (Curtis *et al.*, 1994; Cripps *et al.*, 1995). The protein encoded by the gene has since been shown to be a negative regulator of the cell cycle (Matsumine *et al.*, 1996) however, there have been no publications reporting biallelic inactivation of this gene.

The analyses of polymorphic sites within the MCC gene were included in this study to increase the number of informative cases. In all cases that were informative for both APC and MCC; loss or retention was always seen in both genes i.e. there were no cases showing discordance between the two genes, indicating that this was a minimum area of loss.

To further investigate the role of *APC* in NSCLC tumourigenesis, cases that exhibited LOH at 5q21 were screened for mutations within the *APC* gene. It is unknown whether biallelic inactivation of *APC* is a feature of NSCLC. One study consisting of 7 NSCLC cases that displayed LOH at 5q21 detected no mutations in the remaining *APC* allele (Horri *et al.*, 1992b).

Lung cancer is associated with cigarette smoking, the carcinogen metabolites that are produced commonly form DNA adducts at guanine or adenine nucleotides and analysis of mutations present in tumour suppressor genes and oncogenes in lung cancers typically exhibit a high frequency of mutations at guanine nucleotides (discussed in section 1.1.2, and 1.3). One may therefore expect to see a high level of G-T transversions or A-T transversion mutations in lung cancer cases from individuals who smoked. Throughout the *APC* gene, there are numerous codons that may give rise to stop codons following the production of DNA adducts at a guanine or an adenine residue. One may predict that these codons could become mutation hotspots in lung cancer arising in smokers.

The coding region of the *APC* gene is 8529 nucleotides, to sequence the whole length of the gene for mutations would be extremely expensive and time consuming. Therefore, in

this study I decided to analyse the MCR of the gene as the importance of this region in tumour suppression had been reflected by previous mutation studies in other neoplasms. The method used to screen this region for mutations was single strand conformational polymorphism analysis (SSCP), a relatively quick and simple screening method. No mutations were detected in the 30 cases analysed. This negative result may raise questions on the sensitivity of this method despite all colorectal cancer cases with known mutations, which were used as positive controls, being detected by this method. SSCP analysis has been shown to detect >90% of single base substitutions in PCR fragments of 200 bp and >80% in fragments of 400 bp in size (Suzuki *et al.*, 1990; Gaidano *et al.*, 1991; Hayashi 1991; Hayashi and Yandell 1993). In this study the MCR was amplified by PCR in fragments ranging from 200-300 bp (see section 3.3.4.1), therefore one may assume a detection rate of between 80-90%. Other than the size of the fragment, one must consider that the position of the mutation may influence whether the mutation is detected since sequences neighbouring the mutation have been shown to affect conformation and hence mobility within the gel (Sheffield *et al.*, 1993; Glavac and Dean 1993). Another consideration is the presence of normal DNA present within the extracted tumour DNA samples. In the case of normal DNA contamination, detection of the mutation within the tumour DNA will depend on the degree of separation that can be achieved between wildtype and mutant bands. Therefore, in this study samples were electrophoresed to the bottom of the gel to ensure maximum separation .

Despite the possible false negatives that may occur using this method, I feel that from the absence of mutations within these 30 cases one can say that mutations within the MCR of the *APC* gene are rare.

Immunohistochemistry was then used to detect mutations throughout the whole of the *APC* gene. Tissue sections were screened for full-length protein by immunohistochemistry using an antibody raised to the carboxyl terminus of the APC protein. This method also has its limitations with only mutations leading to the truncation of the APC protein being detectable.

Immunohistochemical studies had previously reported that APC protein is ubiquitously

expressed within the lung (Midgley *et al.*, 1997). The analysis of normal lung tissue in this study confirmed these results with a diffuse cytoplasmic staining pattern being noted in the majority of cells. Interestingly, previous reports have indicted that *APC* expression is restricted to regions where cell replication has ceased and terminal differentiation is established (Miyashiro *et al.*, 1995; Midgley *et al.*, 1997), and within the lung mitotic index is very low. Relatively high levels of protein were detected within the epithelial cells lining the bronchioles with accentuated staining at the lateral and apical surfaces. The lateral membranes are typically sites for adhesion junctions; junctions that are required not only for intercellular adhesion, but also for intercellular signalling and cell polarity. APC has been shown to co-localise with both α- and β-catenin at the lateral membranes of colonic epithelium, therefore it would be interesting to confirm whether APC co-localises with these proteins within the lung. Apical staining has also been noted in colonic epithelium and co-localisation studies have shown the absence of α- and β-catenin but co-localisation with actively growing microtubules, suggesting APC plays a role in cell migration (discussed in section 2.3.4). Again, studies are needed to explore the co-localisation of APC with other proteins within the lung.

In some cells, the APC protein was detected in the nucleus. The nuclear localisation of APC has been previously reported using immunohistochemical techniques and confirmed by subfractionation experiments (Neufeld and White, 1997). Two putative nuclear localisation sequences have been identified within APC, these sequences begin at amino acids 1773 and 2054. The significance of nuclear localisation of APC has not been determined.

Full length APC protein was detected in all NSCLC cases that exhibited LOH at 5q21. As seen in normal lung tissue, APC protein was detected in both the cytoplasm and in some cells within the nucleus. These results reveal that the remaining *APC* allele is producing full length APC protein. Thirty cases were analysed, therefore one could state that mutations that result in the production of a truncated APC protein are rare in ADC and SCC of the lung. The detection of full length APC in these cases also suggests that epigenetic factors do not lead to loss of *APC* expression in these cases. Hypermethylation

of the *APC* promoter has previously been reported to occur in patients with colorectal carcinoma and it is thought to be involved in the progression of colorectal cancer (Hiltunen *et al.*, 1997).

According to Knudson's hypothesis tumour suppressor genes act recessively at the cellular level so that both copies of the gene must be inactivated for growth suppressive functions to be eliminated (Knudson 1978). LOH has been detected in a proportion of NSCLC cases but in all cases showing LOH, no mutations have been detected within the remaining *APC* allele and full length protein is produced. The methods used in this research have not eliminated the possibility of mutations outside the MCR (the remaining 1/3 of the gene) that do not result in the production of a truncated protein product. These mutations may be detected by sequencing the remaining portion of the *APC* gene or alternatively carrying out protein function tests such as electrophoresis mobility shift assays (EMSAs) however, both these methods are expensive and time consuming.

Considering all the results within this study and research that has examined the *APC* gene in other types of cancers, I feel that it is unlikely that the *APC* gene is inactivated in ADC and SCC of the lung.

Failure to detect mutations within the *APC* gene may be a reflection of mutations in other components of the wnt signalling pathway. The presence of mutations within other components of the wnt signalling pathway may also explain the overexpression of two wnt target genes, *CCND1 (Cyclin D1)* and c-*MYC, in NSCLC* (see sections 1.3.1.2b, 1.3.4.2, 2.6.10.5 and 2.6.10.6). Overexpression of cyclin D1 has been reported to occur in approximately 50% of NSCLC cases, of which amplification of the *CCND1* accounts for around 2-30% cases, whilst *c-MYC* is overexpressed in 50% of cases and has been shown to be amplified in 10% of these cases. In the remaining cases, the mechanism of gene deregulation is unknown. Both *CCND1* and *c-MYC* are also activated by pathways controlled by *K-RAS* (MEK/ERK and PI3K), therefore gene expression may be deregulated due to activating mutations in *RAS* or by deregulation of any other components of these pathways (see section 1.3.1.1).

If mutations in components of the wnt signalling pathway, other than *APC*, are present

then subsequent mutations within the *APC* gene may not give rise to an increased growth advantage. This has been noted in colorectal cancer where cases have only been shown to harbour a mutation in the *β-catenin* gene or *APC* but never both genes (Sparks *et al.*, 1998). To date no publications have reported the status of the genes *WNT, frizzled, dishevelled, Axin* or *β-catenin* in lung cancer.

If the *APC* gene did play a role in lung carcinogenesis, one would expect individuals with germline mutations in the *APC* gene to be at increased risk of lung cancer, this is not the case.

LOH at 5q21 is clearly a relatively common and relatively early event in pulmonary ADC and SCC tumourigenesis. The data presented here suggest that it is unlikely that APC plays a role in these cancers, therefore investigations should now concentrate on the identification and characterisation of other putative tumour suppressor genes within the 5q21 region. The *MCC* gene was initially considered as a putative tumour suppressor gene although further studies have not supported this theory. Other known genes within this region include the *FER*, a tyrosine kinase homologous to *SRC*, *TB1* which shares similarities with ADP, ATP carrier/ translocator protein family, *TB2* with no similarities with other known genes; and the *SRP19* gene (Joslyn *et al.*, 1991; Kinzler *et al.*, 1991; Kinzler *et al* 1991). These genes are currently being fully characterised, to date no mutations within these genes have been reported.

Chapter 4: *K-RAS* mutations in adenocarcinoma and alveolar atypical hyperplastic lesions of the lung

4.1 Introduction

Multiple genetic alterations are involved in lung tumourigenesis and include the activation of oncogenes and inactivation of tumour suppressor genes. The proto-oncogene *K-RAS*, which encodes a protein that functions as a guanosine diphosphate /guanosine triphosphate (GDP/GTP) regulated switch transducing extracellular stimuli to cytoplasmic signal transduction cascades, has been shown to be frequently mutated in 25-50% of NSCLC.

The majority of publications report that *K-RAS* mutations occur primarily in the ADC histological subtype. The reported frequency varies between 20-50% of tumours, whilst in SCC mutations are infrequent (0-13% of cases) (Rodenhuis *et al.*, 1988; Slebos *et al.*, 1989; Reynolds *et al.*, 1991, Mitsudomi *et al.*, 1991; Slebos *et al.*, 1991a; Mills *et al.*, 1995; Keohavong *et al.*, 1996; Graziano *et al.*, 1999; Gealy *et al.*, 1999). Conversely, some studies have reported that mutations occur in SCC with a similar frequency to ADC (Rosell *et al.*, 1993; 1996; Gao *et al.*, 1997). It is currently unclear why SCC are mutated in some cohorts but not others. Studies have analysed a full spectrum of SCC; different tumour stages and SCC obtained from smokers and non-smokers. One may suggest that variation in frequency may be a reflection of geographical differences.

The majority of *K-RAS* mutations detected in pulmonary ADC and SCC occur at codon 12 (80%), with a lower frequency occurring at codons 13 and 61. Mutations at these sites result in the production of a protein product that remains in the active GTP-bound state, which results in constitutive signalling leading to an oncogenic phenotype (discussed in detail in section 1.3.1.1b). The majority of studies report that the mutation spectrum seen at codon 12 in ADC and SCC, consists of G-T transversions at position 1 (60% of cases), G-A transition at position 2 (20% of cases) and G-T transversion at position 2 (15% of cases) (Rodenhuis *et al.*, 1987; Rodenhuis *et al.*, 1988; Reynolds *et al.*, 1992; Slebos and

Rodenhuis, 1992; Vachtenheim *et al.*, 1995). One study does differ and reports that within SCC, the majority of mutations at codon 12 position 1 and position 3 are A-T transversion (Gao *et al.*, 1997). The mutation spectrum seen in the *K-RAS* gene is consistent with the formation of DNA adducts at guanine or adenosine residues. These adducts can result from the carcinogenic properties of tobacco smoke, and as expected, the presence of *K-RAS* mutations has been significantly associated with tobacco exposure (see section 1.1.2; Rodenhuis *et al.*, 1988; 1992; Slebos *et al.*, 1991b; Westra *et al.*, 1993; Gealy *et al.*, 1999 and references therein). Interestingly, a high frequency of G-A transitions has been noted in non-smokers with Goa and co-workers (1997) reporting that 11/11 tumours exhibited this mutation.

A spectrum of well-defined histological lesions have been associated with the development of several epithelial tumours. The analysis of these lesions have aided the understanding of the molecular events underlying tumourigenesis. In pulmonary SCC the progressive morphological changes include goblet cell hyperplasia, basal cell hyperplasia, squamous cell metaplasia, atypia or dysplasia, carcinoma in situ (CIS) and invasive cancer (Saccomanno *et al.*, 1974). Wistuba and co-workers (1999) have recently carried out an extensive analysis of the molecular abnormalities involved in the pathogenesis of the multistep development of pulmonary SCC. The investigators showed that there was an increase in the frequency of LOH with increasing severity of histopathological changes. The earliest and most frequent regions of allelic loss occurred at 3p21, 3p22-24 3p25 and 9p21 (reviewed in section 1.3.3.2). Loss of heterozygosity within the *p53* gene was seen dysplastic lesions and CIS, suggesting that *p53* mutation is a relatively late event in SCC tumourigenesis (see section 1.3.2.2).

Even though many genetic changes have been identified in pulmonary ADC, little is known about the chronology of their development due to the absence of well-characterised premalignant lesions. Furthermore, as ADC differs significantly in their histology, it is likely that different premalignant lesions exist. Areas of alveolar atypical hyperplasia (AAH) are considered as potential premalignant lesions of ADC. These lesions arise in the periphery of the lung (within the alveolar compartment) and are

characterised by progressive cellular atypia of type II pneumocytes (discussed in detail in section 1.2.2).

Immunohistochemical analysis of AAH lesions have shown that these lesions overexpress the tyrosine kinase receptor *ERB-B2* (see section 1.3.1.3.a), exhibit high levels of p53 protein (see section 1.3.2.2) and carcinoembryonic antigen (Hamanda *et al.*, 1995 Nakanishi, 1990; Carey *et al.*, 1992; Kitamura *et al.*, 1996). Aberrant expression of these proteins has been reported in pulmonary ADC, therefore one may suggest that aberrant expression of these genes is an early event in ADC tumourigenesis with AAH lesions representing preneoplastic lesions.

Oncogenic activation of *K-RAS* has been reported to be an early event in the multistep process of colorectal and pancreatic cancer. *K-RAS* mutations have been detected in colorectal adenomas and in aberrant crypt foci (possible precursor lesions of colorectal ADC) (Pretlow 1993; Yamashita *et al.*, 1995; Losi *et al.*, 1996; Shivapurkar *et al.*, 1997). Within the pancreas mutations have been detected in small pancreatic adenomas (Schaeffer 1994) and in atypical hyperplastic lesions and non-neoplastic (hyperplastic and metaplastic) ductal cells of patients with pancreatic cancer (Tabata *et al.*, 1993; Sugio *et al.*, 1997; Matsubayashi *et al.*, 1999; Luttges *et al.*, 1999).

Investigations have shown that *K-RAS* codon 12 mutations are present in foci of non-invasive carcinoma (CIS) and are homogeneously distributed throughout small ADC (Li *et al.*,1994a; Sugio *et al.*, 1994). These results suggest that *K-RAS* mutations are an early event occurring prior to clonal expansion of the tumour and due to the absence of recognised precursor lesions it has been difficult to further define the timing of *K-RAS* mutations. Investigations have reported that there is no significant difference between the frequency of *K-RAS* mutations and tumour stage, this suggests that *K-RAS* mutations are not associated with tumour progression (Sugio *et al.*, 1992; Rodenhuis 1992).

4.2 Aims

The research described in this chapter aimed to determine firstly, the frequency of *K-RAS* codon 12 mutation in a series of ADC cases. Secondly, determine whether there was an association between *K-RAS* mutations and the site of origin of the ADC; a comparison between parenchymal ADC arising in the distal parenchyma and bronchial ADC taking origin (like squamous and small cell tumours) from the bronchi. Finally, AAH lesions were examined to identify whether these lesion harbour *K-RAS* mutations.

4.3 Materials and methods

4.3.1 Collection of ADC cases and extraction of DNA

Collection of ADC and extraction of DNA has previously been described within section 3.1.1.

4.3.2 Identification of alveolar atypical hyperplastic lesions and extraction of DNA

Alveolar atypical hyperplastic lesions were identified in non-tumour parenchyma from patients previously diagnosed with ADC and also patients who had been biopsied for possible lung cancer. Ten-micron sections from biopsies containing areas of AAH were cut and de-paraffinised as described in section 3.3.2. To visualise areas of AAH, sections were incubated for 3 minutes in eosin and then rinsed twice in sterile DDW. Slides were examined under the microscope and areas of AAH marked using a permanent marker pen on the underside of the slide. The area of AAH was isolated by gently scraping away surrounding tissue using a sterile scalpel then pipetting 2-5µl of DNA lysis buffer (Appendix A) onto the lesion and drawing up the tissue in the buffer. The buffer was then pipetted into a sterile eppendorf containing 50-100µl of lysis buffer, the volume dependent on the size of the lesion. Proteinase K (20 mg/ml) was then added to give a final concentration of 1.25µg/µl. The samples were incubated for a period of 1 hour at 55°C, gently vortexed and incubated for a further 3 hours. Following incubation, sterile EDTA, pH 8.0 (100 mM stock solution) was added to a final concentration of 1 mM to prevent bacterial growth in the sample and prevent digestion of DNA by nucleases. DNA samples were aliquoted into 50µl samples and frozen at –20°C until required.

4.3.3 Detection of *K-RAS* codon 12 mutations

PCR was applied to tumour and normal DNA extracted from ADC and AAH lesions to amplify a 157 base pair product encompassing codon 12 of the *K-RAS* gene. The primers used have been described previously and incorporate nucleotide changes designed to

create two *BstN1* restriction endonuclease sites (CCTGG) in the amplified sequence (Jiang *et al.*, 1989).

Primer sequences were:

Upstream 5' ACTGAATATAAACTTGTGGTAGTTGGA**C**CT 3' within exon 1.

Downstream primer 5' TCAAAGAATGGTCCTGGACC 3' within intron 1.

The exon 1 primer was degenerate, substituting guanine with cytosine (bold) at the first position of codon 11, creating a *BstN1* restriction endonuclease site (underlined) that extends into position 1 and 2 of codon 12 (outside the primer sequence). Mutations in either of the first two positions of codon 12 result in loss of this restriction site, and following incubation with *BstN1* can be identified by gel electrophoresis. The downstream intron 1 primer contained a *BstN1* restriction endonuclease site that acted as a cleavage control, cleaving a 14 bp fragment. Figure 13 shows the primer locations and the possible PCR product fragment sizes that are seen in the case of mutation or no mutation within position 1 or 2 of codon 12 following incubation with *BstN1*.

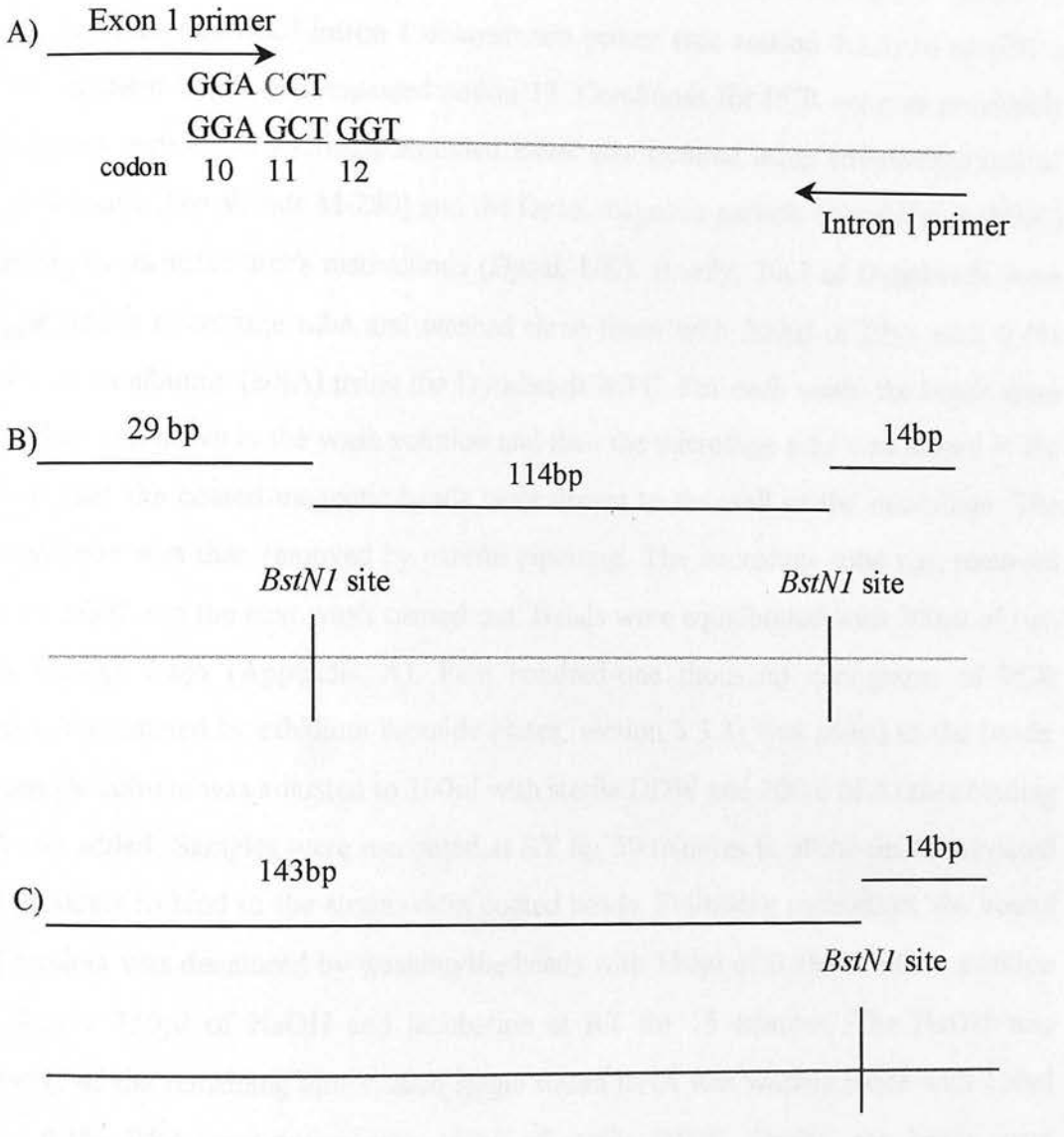
PCR reactions were prepared as described in section 3.3.4.1, and subject to PCR cycling parameters of 94°C for 5 minutes followed by 34 cycles of 94°C for 30 seconds to denature, 55°C for 30 seconds to anneal and 72°C for 30 seconds for extension. A final extension time of 10 minutes at 72°C was also included. To ensure amplification of the fragment had been successful, the PCR products were electrophoresed through a 3% agarose gel containing 0.1 mg/ml of ethidium bromide in 1 x TBE running buffer and visualised under ultra-violet light (described in section 3.3.4.2). Approximately 0.2µg of PCR product was added to 2 units of *BstN1* restriction endonuclease and 2µl of restriction enzyme buffer. The volume was adjusted using sterile DDW to 20µl. Endonuclease reaction mixes were incubated for 2 hours at 60°C. Following incubation samples were again analysed on a 3% agarose gel and the size of the cleaved products determined using a molecular weight marker V (Boehringer Manheim).

Figure 13: Detection of mutations within codon 12 of the *K-RAS* gene

A PCR fragment of 157 bp, which encompassed codon 12 of the *K-RAS* gene, was amplified using PCR. One primer was situated within exon 1 (upstream primer) and contained a degenerate nucleotide, a cytosine substitution corresponding to the first position of codon 11, thus creating a *BstN1* recognition site (CCTGG) extending into position 1 and 2 of codon 12, nucleotide positions outside the primer sequence. The *BstN1* restriction site is only present if there are no mutations within position 1 or 2 of codon 12. The down stream primer was situated within intron 1 and also contained a *BstN1* site, this primer was not degenerate. Primers are shown in figure 13 (a).

Following amplification of the 157 bp fragment, PCR products were incubated with the restriction endonuclease *BstN1*. In the case of no mutations within codon 12, the PCR product yielded 3 fragments of 114 bp, 29 bp and 14 bp, as exemplified in figure 13 (b). If a mutation was present in position 1 or 2 of codon 12 the upstream *BstN1* restriction site is lost (29 bp fragment) therefore yielding protein products of 143 bp and 14 bp, exemplified in figure 13 (c).

Figure 13: Detection of mutations within codon 12 of the *K-RAS* gene



B) ——— DNA with no codon 12 mutation at position 1 or 2

C) ——— DNA with a codon 12 mutation at position 1 or 2

4.3.4 Sequencing codon 12 of the *K-RAS* gene

4.3.4.1 Preparation of single stranded PCR product using Dynabeads

DNA from samples shown to carry codon 12 mutations by *BstN1* restriction analysis were sequenced to determine nucleotide change. DNA samples were re-amplified by PCR using a biotinylated upstream primer 5' GACTGAATATAAACTTGTGG 3' (exon 1) and the previous described intron 1 downstream primer (see section 4.3.3) to amplify a 145 bp fragment which encompassed codon 12. Conditions for PCR were as previously described in section 4.3.3. Single stranded DNA was isolated using streptavidin coated magnetic beads (Dynabeads M-280) and the Dynal magnetic particle concentrator (MPC) according to manufacturer's instructions (Dynal, UK). Briefly, 30µl of Dynabeads were pipetted into a microfuge tube and washed three times with 300µl of PBS with 0.1% bovine serum albumin (BSA) using the Dynabeads MPC. For each wash, the beads were pipetted up and down in the wash solution and then the microfuge tube was placed in the MPC so that the coated magnetic beads were drawn to the wall of the microfuge. The wash solution was then removed by careful pipetting. The microfuge tube was removed from the MPC and the next wash carried out. Beads were equilibrated with 300µl of two times binding wash (Appendix A). Five hundred-one thousand nanograms of PCR product (determined by ethidium bromide plates, section 3.3.3) was added to the beads. The sample volume was adjusted to 100µl with sterile DDW and 100µl of 2 times binding wash was added. Samples were incubated at RT for 30 minutes to allow the biotinylated PCR products to bind to the streptavidin coated beads. Following incubation, the bound PCR product was denatured by washing the beads with 150µl of 0.15M NaOH, addition of a further 150µl of NaOH and incubation at RT for 15 minutes. The NaOH was removed and the remaining biotinylated single strand DNA was washed twice with 150µl PBS / 0.1% BSA, and twice with 150µl of sterile DDW. Finally, the beads were resuspended in 7µl of sterile DDW and the DNA template stored at 4°C until required.

4.3.4.2 Sequencing PCR product

DNA sequencing was carried out using the Sequenase Version II sequencing kit

(Amersham, UK). The sequencing protocol is based on the chain termination method originally described by Sanger *et al.* (1977). To the 7µl of template, previously prepared using Dynal beads, 1µl (1 pM) of sequencing primer, namely the downstream intron 1 primer, and 2µl of reaction buffer (Appendix A) were added. Reactions were placed in a beaker of water at 65°C, which was left to cool on the bench to 35°C over a period of 15 to 30 minutes. This allowed the primer to anneal to the template. Reactions were microfuged briefly and placed on ice. One hundred millimolar dithiothreitol, 2µl labelling mix (Appendix A), 5µCi (^{35}S) dATP and 2µl (3.25 units) of Sequenase enzyme were added. Samples were vortexed gently and the labelling reaction was carried out at RT (24-26°C) for 5 minutes. Termination of the sequencing reaction was carried out by the addition of 3.5µl of labelling reaction to 2.5µl of each of the dideoxynucleotides (ddATP, ddCTP, ddGTP, ddTTP) which had been preheated to 37°C, the termination mix was incubated for 5 minutes at 37°C. Four microlitres of sequencing stop solution (Appendix A) was added to the termination reactions following the 5 minute incubation.

Sample were denatured by incubating at 95°C for 5 minutes then 3µl of the sample was loaded onto a 6% polyacrylamide, 8M urea sequencing gel (Appendix A) which had been pre-run to warm to 50°C. The samples were electrophoresed at 70 W, for 2 hours in 1 x TBE running buffer. Following electrophoresis gels were fixed for 5 minutes in 10% (v/v) methanol / 10% (v/v) glacial acetic acid. Fixed gels were blotted onto Whatman 3MM paper and dried under vacuum at 120°C for 2 hours, then exposed to Kodak Biomax MR autoradiograph film at RT within a radioactive cassette for 2-3 days. Autoradioactive film was subsequently developed using an automated film developer and the DNA sequence read.

4.3.5 Conformation of G-T transversion at position 1 and 2 within an alveolar atypical hyperplastic lesion

Following DNA sequencing one AAH lesion was shown to harbour two nucleotide

substitutions, G-T transversions at position 1 and 2 of codon 12 (see results section 4.4.2). To determine whether this mutation was present within an individual cell or was a result of a mixed population of cells, the PCR product was cloned and then the cloned product sequenced. Cloning was carried out using the Original TA Cloning Kit (Invitrogen). This kit was selected as amplified PCR products could be cloned directly into the pCR2.1 plasmid vector supplied with the kit without need to enzymatically modify the PCR product. This is possible if Taq polymerase is used to amplify the product. Taq polymerase has a non template-dependent activity that adds a single deoxyadenosine (A) to the 3' ends of the PCR products, the linearised plasmid vector pCR2.1 has single 3' deoxythymidine (T) residues, which provide complementary ends for ligation with the PCR product.

The pCR2.1 plasmid vector contains the ampicillin resistance gene, therefore following transformation into competent cells, in this case INVαF' cells which do not contain ampicillin resistance (Invitrogen), only cells which contain the vector will be able to grow when plated onto Luria-Bertani (LB) plates containing the antibiotic. The colonies on these plates may contain the vector with the PCR product inserted or may be self-ligated vector (see below). The colonies growing on the plates can be screened for these two possible outcomes using the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). X-gal is cleaved by β-galactosidase, releasing a blue indolyl derivative. The pCR2.1 vector harbours the β-galactosidase gene, the gene promoter and lacZα fragment of the gene are situated 5' to the linearisation site (site of PCR product insertion) whilst the remaining proportion of the gene is situated 3' to the linearisation site. Following self-ligation, the fragments of the β-galactosidase gene are within a continuous trans-position and the β-galactosidase enzyme is produced. In this case, in the presence of X-gal, colonies turn blue. When the PCR product is inserted into the linearisation site, the β-galactosidase gene is interrupted and β-galactosidase is not produced, therefore no chromogenic substrate is produced in the presence of X-gal and colonies remain white.

4.3.5.1 Ligation reactions

According to manufacture's directions, PCR products which were less than 1 day old (as single 3'A-overhangs on the PCR product degrade with time) were used to set up ligation reactions. Ligation reactions contained PCR product and vector at a molar ratio of 1:1, see section 5.3.1.1.e for formula. Two nanograms of PCR product (as determined by ethidium bromide plates, see section 3.3.3) was mixed with 50 ng of pCR2.1 vector and 1µl of 10x ligation buffer (Appendix A) and 1µl T4 DNA ligase (4.0 Weiss units) was added. The volume of the ligation mix was adjusted to 10µl by the addition of sterile DDW.

A control ligation reaction was also prepared to detect false positive colonies. As seen with the 3' A overhangs present on PCR products the 3'T-overhangs present on the vector may also be lost, this has been shown to be a result of time or due to incorrect storage or repeated freezing and thawing. Loss of the 3'T-overhangs can result in blunt-ended self-ligation of the vector. This can cause a frameshift of the lacZ gene, which prevents or reduces β-galactosidase production, therefore white or light blue colonies that do not contain the insert may be detected. To control for this event, a ligation reaction was set up that contained no PCR product.

Ligation reactions were incubated at 14°C overnight. Ligations were then stored at –20°C prior to transformation.

4.3.5.2 Transformation into INVαF' competent cells

Competent cells were removed from –70°C and thawed on ice. Fifty microlitres of cells were placed in a pre-chilled 15 ml Falcon tubes and β-mercaptoethanol (supplied with the cells) added to give a final concentration of 25 mM. Cells were incubated for 10 minutes on ice then 2µl of ligation mix added and the tubes incubated on ice for a further 30 minutes. Tubes were placed in a 42°C water bath for 45 seconds and then immediately placed on ice for 2 minutes. Nine hundred microlitres of preheated (42°C) SOC medium (Appendix A) was added and the tubes incubated at 37°C for 1 hour shaking in an orbital shaker at 225 rpm.

One hundred microlitres and 200µl of each transformation was spread onto separate LB agar plates containing X-Gal and 50µg/ml of ampicillin (see Appendix A). The transformation solution was left to absorb into the agar plate for approximately 5 minutes and then plates were inverted and placed at 37°C for at least 18 hours. Plates were then placed at 4°C for 2-3 hours to allow development of the colony colour.

4.3.5.3 Bacterial cultures and plasmid isolation

Individual white colonies (clones) from the agar plates were transferred into 3 mls of LB broth supplemented with 50µg/ml ampicillin (Appendix A) and incubated shaking at 225 rpm at 37°C overnight. Plasmid DNA was purified using the Wizard™ plus Minipreps DNA Purification System (Promega UK). The protocol was as follows; 1.5 ml of bacterial culture was centrifuged at 13,000 rpm to pellet the cells. The supernatant was discarded and the cell pellet resuspended in 200µl of cell resuspension solution (Appendix A) and transferred into a 1.5 ml microfuge tube. Two hundred microlitres of cell lysis solution (Appendix A) was added and the suspension mixed by gently inverting the tube. Two hundred microlitres of neutralisation solution (Appendix A) was added and the solution mixed by inverting the tube. The tubes were microfuged at 13,000 rpm for 5 minutes to pellet the cell debris and protein. The cell lysate was then loaded onto a minicolumn/syringe assembly containing 1 ml of resin. The lysate was drawn down the column via vacuum and the column washed with column wash solution (Appendix A). Plasmid DNA was eluted from the minicolumn by the addition of 50µl of sterile DDW, preheated to 65°C, and microfuging at 13,000 rpm for 20 seconds. Plasmid DNA was collected in a sterile 1.5 ml microfuge tube and stored at 4°C.

4.3.5.4 Sequencing plasmid DNA

Plasmid DNA was sequenced using the T7 Sequenase quick denature plasmid sequencing kit (Amersham Life Sciences). The sequencing protocol is based on the chain termination method originally described by Sanger *et al.* (1977) and has been previously described in section 4.3.3. The production of single stranded DNA differs in this protocol, in the

previous case Dynal beads were used. To obtain single stranded DNA, NaOH was added to 0.5-3μg of plasmid DNA to give a concentration of 0.2M, the reaction mix was adjusted to 8μl with the addition of sterile DDW. Two picomole of sequencing primer was added to each sample and left to anneal to the template DNA for 10 minutes at 37°C, then placed on ice. The sample was neutralised by the addition of HCl to a final concentration of 0.2M. Two microlitres of plasmid reaction buffer (Appendix A) was added and samples left to anneal at 37°C for a further 10 minutes then chilled on ice. One hundred millimoles of dithiothreitol, labelling mix (Appendix A), 5μCi (^{35}S) dATP and Sequenase enzyme were added. Samples were vortexed gently and labelled reactions were carried out at RT for 5 minutes. Three and a half microlitres of the labelling reaction was then added to 2.5μl of each of the dideoxynucleotides at 37°C for 5 minutes. The reaction was terminated by addition of 4μl of sequencing stop solution (Appendix A). Samples were loaded onto a DNA sequencing gel and electrophoresed using conditions described in section 4.3.4.2. Sequencing gels were subsequently dried and exposed to x-autoradiograph film as described in section 4.3.4.2.

4.4 Results

4.4.1 Mutations at codon 12 of the *K-RAS* gene in adenocarcinomas

Mutations in codon 12 of the *K-RAS* oncogene were detected using a combination of PCR amplification and restriction endonuclease cleavage analysis (Jiang *et al.*, 1989). This method is highly specific, with loss of a *BstN1* restriction endonuclease recognition site being diagnostic for the presence of mutations at position 1 or 2 of codon 12. An additional *BstN1* restriction endonuclease recognition site offers a control for the restriction endonuclease cleavage .

Sixty-five ADC samples were assayed for the presence of mutations in codon 12 (Data are shown in Appendix D). Mutations in codon 12 were detected in 16 cases (25%). There was no statistically significant difference between the presence of *K-RAS* mutation and tumour stage, Fisher Exact Test $p>0.2$ in all cases (see table 8a). Examples of normal and tumour DNA pairs analysed for mutations in codon 12 of the *K-RAS* gene by restriction endonuclease cleavage are shown in figure 14 (a).

DNA from tumours showing mutation were sequenced. Nine tumours harboured G-T transversions at position 1 of codon 12, 6 tumours G-T transversions at position 2 and one tumour a G-A transition at position 2. Sequencing gels from two cases with mutations at codon 12 are shown in figure 14 (b)(c).

Sub-classification of ADC into bronchial and parenchymal origin revealed that 9/26 (35%) of parenchymal ADC contained mutations in position 1 or 2 of codon 12 whilst 0/12 bronchial ADC were mutated at these positions, see table 8b. This difference was shown to be statistically significant by Fisher's exact test (two-tailed test; $p=0.0355$).

There was no statistically significant difference between the presence of *K-RAS* codon 12 mutation and tumour stage in parenchymal subtypes, as determined by Fisher's exact test (two-tailed test;$p>0.1$).

Twenty seven of the 65 ADC cases screened for mutations could not be classified into either of the two subgroups, of these 6 (22%) contained a mutation at codon 12.

Table 8a: Analysis of *K-RAS* codon 12 mutations with tumour stage in ADC

Tumour Stage	No. of tumours with *K-RAS* mutations/total number of tumours
Stage 1	11/38
Stage II	3/15
Stage 111	2/12

Statistical analysis shows there is no statistically significant difference between the presence of *K-RAS* codon 12 mutation and different tumour stages as determined by Fisher's exact test (two-tailed; $p>0.2$ in all cases).

Table 8b: Analysis of the presence of *K-RAS* codon 12 mutations in relation to origin of ADC

Origin	Mutation	No mutation
Bronchial	0	12
parenchymal	9	17

Statistical analysis shows that there is a statistically significant difference between the presence of *K-RAS* codon 12 mutation and the origin of the ADC. Fisher's exact test, two-tailed analysis, $p=0.0355$.

Table 8c: Analysis of *K-RAS* codon 12 mutation and tumour stage in parenchymal ADC

Tumour stage	No. of parenchymal ADC with *K-RAS* mutations/ total number of parenchymal ADC
Stage I	7/14
Stage II	2/9
Stage III	1/3

There was no statistically significant difference between the presence of *K-RAS* codon 12 mutation with different tumour stages in the parenchymal subtypes, as determined by exact test (two-tailed test; $p>0.1$ in all cases).

Figure 14: Mutational analysis of *K-RAS* codon 12

Figure 14 a) Paired tumour (T) and normal (N) DNA samples were subjected to PCR to amplify a fragment of the *K-RAS* gene encompassing codon 12. PCR products were incubated with the restriction endonuclease *BstN1* and then electrophoresed through a 3% agarose gel. PCR product not subjected to restriction endonuclease cleavage and molecular weight markers were also loaded onto the gel, lanes U and M respectively.

Case 1 is an example of an ADC which harbours a mutation at position 1 or 2 of codon 12. Restriction endonuclease analysis of the amplified tumour DNA results in cleavage of the 157 bp PCR product to yield 143 bp and 14 bp fragments (the latter fragment is not visible on this gel).

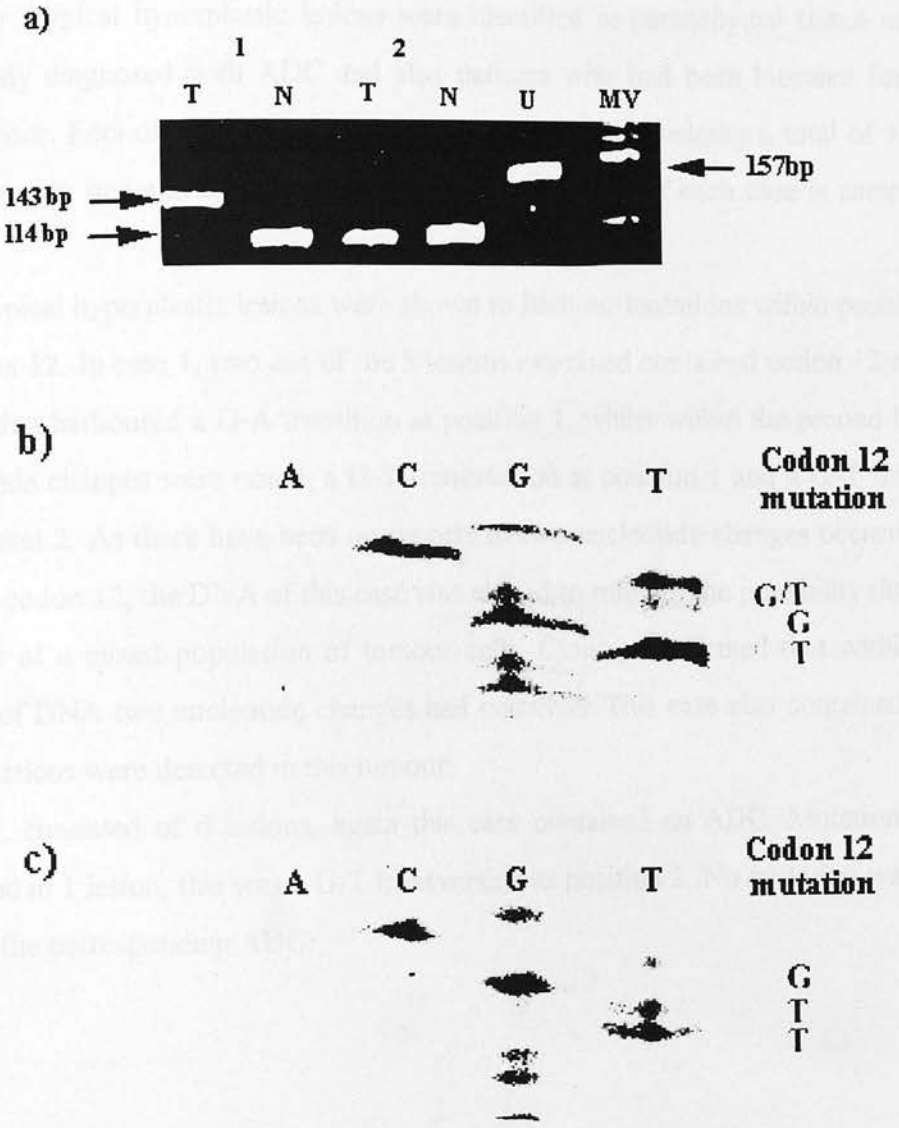
The PCR product amplified from normal DNA of this patient is cleaved to give a 114 bp, 29 bp and 14 bp fragments (the latter two fragments are not visible on this gel), a restriction pattern indicative of no mutations at positions 1 or 2 of codon 12.

Within the tumour sample a faint band is also detected at 114 bp and is a result of incomplete cleavage or contaminating normal DNA.

In case 2, following incubation with *BstN1*, the PCR products of both normal and tumour DNA are cleaved to give 114 bp, 29 bp and 14 bp (the latter two fragments are not detected on this gel). This restriction pattern is indicative of no mutations at position 1 or 2 in codon 12.

Figure 14 b) and c) Following the identification of mutations within codon 12, through the use of PCR and restriction endonuclease cleavage, DNA from cases shown to be mutated were amplified by PCR and the product sequenced. Figure 14b is an example of an ADC, which harbours a G-T transversion at position 1 of codon 12. Figure 14c shows an ADC case with a G-T transversion at position 2 of codon 12. In both examples, the normal DNA base (G) can be seen as a faint band on the gel. This is a result of contaminating normal DNA.

Figure 14: Mutational analysis of codon 12 of the *K-RAS* gene



4.4.2 Mutations at codon 12 of the *K-RAS* gene in alveolar atypical hyperplastic lesions

Alveolar atypical hyperplastic lesions were identified in parenchymal tissue of patients previously diagnosed with ADC and also patients who had been biopsied for possible lung cancer. Foci of AAH were available from 16 patients yielding a total of 32 lesions, with between one and six distinct foci per case. Analysis of each case is summarised in table 9.

Five atypical hyperplastic lesions were shown to harbour mutations within position 1 or 2 of codon 12. In case 1, two out of the 5 lesions examined contained codon 12 mutations. One lesion harboured a G-A transition at position 1, whilst within the second lesion two nucleotide changes were noted; a G-T transversion at position 1 and a G-T transversion at position 2. As there have been no reports of two nucleotide changes occurring within *K-RAS* codon 12, the DNA of this case was cloned to rule out the possibility that this was a result of a mixed population of tumour cells. Cloning confirmed that within a single strand of DNA two nucleotide changes had occurred. This case also contained an ADC, no mutations were detected in this tumour.

Case 7, consisted of 6 lesions, again this case contained an ADC. Mutation was only detected in 1 lesion, this was a G-T transversion at position 2. No mutation was detected within the corresponding ADC.

Table 9:*K-RAS* mutational analysis in ADC and alveolar atypical hyperplastic lesions.

Case No.	Lesion	*K-RAS* mutation	Codon 12
1	ADC	-	~
	AAH	+	G-A transition pos.1
	AAH	+	G-T transversion pos. 1+2
	AAH	-	~
	AAH	-	~
	AAH	-	~
2	AAH	-	~
	AAH	-	~
3	AAH	-	~
	AAH	-	~
	AAH	-	~
4	ADC	-	~
	AAH	-	~
	AAH	-	~
5	ADC	+	G-T transversion pos.1
	AAH	-	~
6	ADC	+	G-T transversion pos.1
	AAH	-	~
	AAH	-	~
	AAH	-	~
7	ADC	-	~
	AAH	-	~
	AAH	+	G-T transversion pos.2
	AAH	-	~
	AAH	-	~
	AAH	-	~
	AAH	-	~
8	ADC	-	~
	AAH	-	~
9	ADC	+	G-T transversion pos.1
	AAH	-	~
10	ADC	-	~
	AAH	-	~
	AAH	-	~
11	ADC		~
	AAH	-	~
12	ADC	-	~
	AAH	-	~
13	AAH	+	G-T transversion pos.1
14	AAH	-	~
15	AAH	+	G-T transversion pos. 1
16	AAH	-	~

Abbreviations :AAH – alveolar atypical hyperplasia; ADC- adenocarcinoma;

Pos.- position

4.5 Discussion

A retrospective analysis of a series of pulmonary ADC cases revealed that 25% (16/65) of tumours harboured a mutation at either position 1 or 2 of codon 12 of the *K-RAS* gene. This mutation frequency is similar to previous published studies, which have reported a mutation frequency of between 20-50% (see section 1.3.1.1g and 4.1). It is currently unclear as to why the reported frequency of mutation varies so significantly between cohorts. Mutations within the *K-RAS* gene is thought to be a relatively early event, occurring prior to clonal expansion and the presence of *K-RAS* mutation does not increase significantly with tumour stage (Sugio *et al.*, 1992; Rodenhuis *et al.*, 1992; Li *et al.*, 1994a; Sugio *et al.*, 1994). It is therefore unlikely that the number of different stage tumours within each cohort is a governing factor in frequency variation. Analysis of this series of ADC also revealed no association with tumour stage (see Table 8a, Fisher's exact test, two-tailed test; $p>0.2$ in all cases).

Ninety-four percent (15/16) of mutations were G-T transversions, 56% (9/16) occurred at position 1 of codon 12 and 38% (6/16) at position 2. The only other mutation to be detected was a single G-A transition at position 2 (frequency of 6%). G-T transversion have been reported by numerous studies to be the most common mutation within the *K-RAS* gene in pulmonary ADC, occurring in approximately 80% of cases (see section 4.1). Guanine to thymine transversions are typical of the production of DNA adducts at guanine residues. The formation of DNA adducts results in the failure of the DNA polymerases to identify the base and during DNA replication, in the majority of cases, the guanine is substituted for a thymine residue in the daughter strand (see section 1.1.2, 1.3.1.1.g and 2.4). The production of DNA adducts at guanine residues has been shown to be associated with the carcinogens derived from cigarette smoking (see section 1.1.2). In accordance with this, numerous studies have reported that G-T transversions in *K-RAS* occur more frequently in pulmonary ADC of smokers than non-smokers (Rodenhuis *et al.*, 1988; 1992; Slebos *et al.*, 1991; Reynolds *et al.*, 1991; Westra *et al.*, 1993; Gealy *et*

al., 1999 and references therein). Furthermore, G-T transversions are commonly seen in pulmonary ADC arising in mice which have been treated with constituents of tobacco smoke (You *et al.*, 1989; Belinsky *et al.*, 1997; Nesnow *et al.*, 1998). Information regarding the patients smoking habits was not available for this study, but it is likely that ADC from non-smokers were included. If this is the case, this study would suggest that G-T transversions are also a common occurrence in ADC of non-smokers. The presence of G-T transversions in non-smokers may be a result of passive smoking or implicate other environmental pollutants that result in guanine DNA adducts.

The subclassification of ADC into bronchial and parenchymal origin showed that mutations were significantly associated with ADC of parenchymal origin, as determined by Fisher's exact test (two-tailed test; $p=0.0335$) and that within the parenchymal subgroup there was no association between *K-RAS* mutation and tumour stage (Fisher's exact test, two-tailed analysis; $p>0.1$). This is the first reported genetic difference between these two histological subtypes and suggests that different genetic pathways may be involved in the development of these tumours. This may not be entirely unexpected as investigations have already shown that the genetic events involved in tumourigenesis of different histological types of pulmonary cancer vary. For example SCC and ADC exhibit different allelotypes, the *K-RAS* oncogene is frequently mutated in ADC but is infrequent in SCC, and the aberrant expression of several oncogenes and tumour suppressor genes are often reported to be associated more frequently with either of these subtypes (reviewed in section 1.3). Parenchymal and bronchial ADC are histologically different, with parenchymal ADC containing cells characteristic of goblet cells, Clara cells, type II pneumocytes and ciliated bronchiolar lining cells and occasionally neuroendocrine cells whilst bronchial ADC contain bronchial epithelial cells or bronchi glands (see section 1.2.1.2). Bronchial ADC are also reported to be more aggressive than pulmonary ADC (Edwards 1987) suggesting that specific mutations may occur in bronchial ADC which promote metastatic spread. It is interesting to note that like bronchial ADC, SCC and SCLC, that also arise in the central portion of the lung and are associated with smoking, are not frequently characterised as having *K-RAS* mutations.

In this study, I have screened positions 1 and 2 of codon 12 of the *K-RAS* gene in a series of ADC to determine the frequency of mutation and have shown that mutations are associated with parenchymal but not bronchial ADC subtypes. This however does not rule out the possibility that mutations are not present at other sites within the gene. Mutations in the *K-RAS* gene have also been reported to occur at codons 13 and 61, with the frequency of mutation at either of these two sites reported to be less than 20% (see sections 1.3.1.1g). It is possible that mutations are present at either of these codons or in fact at other sites that may result in the oncogenic activation of the *K-RAS* gene (discussed in section 1.3.1.1b). Further mutation analysis could be carried out and initial analysis should focus on codons 13 and 61, and may be extended to screening all exons of the genes using a relatively quick and simple method such as SSCP analysis.

Mutations within the *K-RAS* gene are thought to be an early event in the tumourigenesis of pulmonary ADC, however, as preneoplastic lesion have not been identified, it is unclear at exactly what stage these mutations occur. There is evidence to suggest that alveolar atypical hyperplastic lesion is precursor lesion to parenchymal ADC (see section 1.2.2.2). Unlike areas of alveolar hyperplasia, which are identified as areas of proliferation of type 2 pneumocytes forming as a response to lung injury or part of the reparative process of the lung, AAH lesions consist of hyperplastic type II pneumocytes which exhibit atypical cytological features. These lesions are identified in the absence of any inflammatory reaction and tend to be incidental findings following lung resections from patients with ADC. Reports suggest these lesions occur in 5-25% of primary ADC cases, areas of AAH have also been detected in biopsies taken from patients without ADC (Miller, 1990, Weng *et al.*, 1992; Carey *et al.*, 1992).

In pulmonary ADC, *p53* gene mutations are thought to occur prior to clonal expansion of the tumour, therefore if alveolar atypical hyperplastic lesions are precursor lesions to ADC one may expect *p53* mutations to be present. The p53 protein product has been detected in 5-58% of lesions, this figure shows significant variation which is thought to be a result of methodology and antibody used within each study (Kerr *et al* 1994; Kitamura *et al.*, 1996, Cagle *et al.*, 1996; Slebos *et al.*, 1998).Slebos and co-workers

(1998) have recently reported that *p53* gene mutations are an extremely rare occurrence in AAH and that accumulation of p53 is significantly associated with the grade of the lesion, i.e. low-grade, high grade or AAH-like carcinoma, as defined in section 1.2.2.2. In this study Slebos and co-workers carried out immunohistochemical analysis on 37 AAH lesions and paired overt ADC. DNA from all lesions and tumours exhibiting p53 protein accumulation were sequenced, sequencing was limited to the region within exons 5 to exon 9, a region previously shown to encompass the majority of mutations found in several types of cancer (see section 1.3.2.2). Four AAH lesions, identified in 4 patients were shown to exhibit p53 accumulation. Interestingly, these lesions had been graded as AAH-like carcinoma, which may represent a stage of transition from benign to malignant. It is interesting to note that in SCC, mutations within the *p53* gene are also common in the late pre-invasive stage of SCC formation and are thought to be important in driving progression of severe bronchial dysplasia to an invasive SCC (see section 1.3.2.2). In two of the cases shown to harbour an AAH lesion with p53 protein accumulation, protein accumulation was also detected within the carcinoma. DNA sequence analysis showed that only one of the 4 lesions harboured a mutation within the analysed region of *p53*. In this case, the p53 protein has also been detected within the carcinoma. The AHH lesion and the carcinoma both harboured missense mutations within exon 7, however, the mutation within the AAH lesion was an A-G transition at codon 239 and in the carcinoma a C-T transversion at position 250. This result makes it unlikely that AAH lesions are a result of metastatic spread and suggests that AAH lesions may result from field cancerization of the bronchial epithelium (discussed in detail below).

In this study, I aimed to determine whether areas of AAH, putative preneoplastic lesions of parenchymal ADC, contained mutations within codon 12 of the *K-RAS* gene. Previous investigations suggest that *K-RAS* mutations are an early event in ADC tumourigenesis, occurring prior to clonal expansion of the tumour, therefore identification of mutations within the *K-RAS* gene in areas of AAH would add molecular evidence to support their role as preneoplastic lesions.

DNA was extracted from 32 lesions identified in sixteen patients, 10 of these patients

harboured an ADC, and DNA was also extracted from the tumour. Mutations within codon 12 were detected in 5 AAH lesions (16%). Two positive lesions were identified in case 1, which contained 5 lesions plus an ADC. No mutations were detected in the ADC and different mutations were detected within the two AAH lesions (G-A transition pos.1 and a G-T transversion at position 1 and 2). In case 7, one out of 6 AAH lesions harboured a *K-RAS* codon 12 mutation, no mutation was detected in the corresponding ADC. In cases 13 and 15, which consisted of only one lesion and no ADC, mutations were also detected.

From the different genetic backgrounds, one may argue that AAH lesions are independent entities, independent of ADC and of other lesions, and may result due to a phenomenon known as field carcinogenesis. The field cancerization effect was originally described by Slaughter and colleges over 40 years ago, who reported the development of multiple primary head and neck cancers in heavy smokers (Slaughter *et al.*, 1953). Two distinct theories have been proposed to explain the occurrence of multiple primary and secondary tumours and preneoplastic lesions in the aerodigestive epithelium, however, from current data it is unclear which mechanism is involved and it is possible that a combination of both theoretical mechanisms are responsible for lung carcinogenesis. The monoclonal neoplasia theory proposes that the progeny of a single transformed cell can spread to form multiple pre-invasive lesions and tumours. Evidence for this mechanism would include the occurrence of a common mutation within multiple lesions. In support of this phenomenon, analysis of multifocal lesions often present in patients with oesophageal ADC (Barrett's oesophagus) revealed that ADC and synchronous premalignant lesions frequently harboured the same genetic aberrations (Gleeson *et al.*, 1998). Within the lung, identical abnormalities have been found in SCC and adjacent non-malignant epithelium, these mutations include LOH of chromosome 3p (Sundaresan *et al.*, 1992; Hung *et al.*, 1995; Chung *et al.*, 1995; Kohno *et al.*, 1999), and 9p (Kishimoto *et al.*, 1995; Kohono *et al.*, 1999) and point mutations in the *p53* tumour suppressor gene (Sundaresan *et al.*, 1992; Sozzi *et al.*, 1992; Mao *et al.*, 1994; Chung *et al.*, 1995; Franklin *et al.*, 1997). From the analysis of AAH lesions within this study, it is difficult to

determine whether AAH lesions are a result of monoclonal neoplasia. Firstly, in several cases *K-RAS* mutations were only present in one or two lesions. However, one may argue that lesions are a result of spread of progeny from a single progenitor cell and that mutation of the *K-RAS* gene is an additional genetic event, this phenomenon has been referred to as subclonal drift (Nowell, 1976).

The second theory suggests that continued exposure of the aerodigestive epithelium to carcinogens predispose the entire epithelium to develop multiple, independent pre-invasive lesions that can develop into tumours or which may regress, the field cancerization theory. This theory has been supported by numerous studies that have analysed mutations within the *p53* gene in multiple pre-invasive lesions and multiple primary tumours of the lung and head and neck. These studies revealed that lesions and tumours harbour different genetic mutations (Nees *et al.*, 1993; Chung *et al.*, 1993; Noguchi *et al.*, 1993; Sozzi *et al.*, 1995; Lavieille *et al.*, 1998). Within my study, case 1 was shown to harbour two lesions with different *K-RAS* mutations, this may support the field cancerization theory, although it does not rule out the possibility of subclonal drift.

The mutation spectrum seen within AAH lesions consisted of 7 cases of G-T transversions and only one lesion with G-A transition, this spectrum is consistent with mutations detected in parenchymal ADC further strengthening the proposal that AAH lesions are precursors of parenchymal ADC. The high frequency of G-T transversions is also indicative of exposure to cigarette smoke. An unusual double mutation was noted in one lesion, a G-T transversion at position 1 and 2 of codon 12. Initially it was not clear whether different genetic populations were present within the lesion, some cells harbouring a mutation at position 1 and some at position 2. To rule out this possibility, the DNA from the lesion was cloned and the result revealed that mutations at position 1 and position 2 existed within the same DNA strand. As far as I am aware, there are no other reports of two mutations occurring within the *K-RAS* gene. Because of the nature of the oncogenic activation of *K-RAS*, i.e. one mutation with codon 12 leads to constitutive activation of the *K-RAS* gene, and the rarity of such a double mutation, one may assume that the presence of two mutations does not give the cell a proliferative

advantage.

Since the initiation of this work two studies have reported the presence of *K-RAS* mutations within areas of AAH (Ohshima *et al.*, 1994; Westra *et al.*, 1996). Ohshima and co-worker analysed 20 AAH lesions from 6 patients, all of which had undergone a lung resection for ADC. It is likely that in this study lesions identified as areas of AAH were in fact a result of spread of the ADC since it was stated that lesions were generally contiguous with the carcinoma. *K-RAS* mutations were detected in 9 AAH lesions from 2 patients (8 from one patient and 1 from the other patient). These investigators did not identify what mutation had occurred at codon 12.

A second study, by Westra and co-workers, analysed 41 solitary lesions identified in 18 patients for mutations at codon 12. Sixteen lesions (37%) from 11 patients harboured a mutation, this frequency of mutation is similar to that reported to occur in ADC but is much higher than the frequency detected within my study (16%). The mutation spectrum consisted of 62% G-T transversions, 24% G-A transitions and 12% G-C transversions. Different *K-RAS* mutations were detected in AAH lesions within one patient, this may support the theory of field cancerization. This case also confirmed that it is unlikely that AAH lesions are a result of metastatic spread since different *K-RAS* mutations were detected in AAH lesions compared to a synchronous ADC (Westra *et al.*, 1996). Included in this cohort were 2 patients with LCC, one with a SCC and 1 with a carcinoid tumour, of these cases no mutations were detected within tumour tissue. One LCC harboured a G-A transition at position 2 within the only identified AAH lesion, and a G-T transversion was identified in one of 2 lesions present within the patient with SCC.

AAH lesions and ADC analysed in the study by Westra and co-workers were subsequently classified according to the degree of cellular atypia, grades were low grade AAH, high-grade AAH, AAH-like carcinoma and carcinoma (described in section 1.2.2.2). Statistical analysis revealed that there was no significant difference in the frequency of *K-RAS* mutation with increasing lesion grade. This suggests, assuming that AAH lesions are true preneoplastic lesions, mutations within *K-RAS* gene are a very early event in the development of parenchymal ADC.

A recent study has reported LOH at 3p and 9p in AAH lesions (Kohno *et al.*, 1999).In this study Kohono and co-workers analysed 20 AAH lesions identified in 26 patients, all of which harboured a synchronous ADC. Mutational analysis of the *p53* gene was also carried out. Genetic events were detected in three cases. Two cases exhibited LOH at 3p, in the single identified AAH lesion, LOH was present in the corresponding ADC in both cases. In the third case, consisting of an ADC and a single AAH lesion, a mutation within exon 8 of *p53* was detected in both the AAH lesion and ADC, the mutational event was not identified. This case also exhibited LOH at 9p in the AAH and ADC. The frequency of allelic loss at 3p and 9p was extremely low, 10% and 5% of AAH lesions respectively, compared to that seen in ADC 40% and 36%. This may suggest that LOH at these sites is a relatively late event in ADC tumourigenesis, however as the lesions within this study were not graded no conclusion can be made.

Previous studies have determined that mutations within the *p53* gene are a relatively late event in the pathway of AAH through to overt ADC. To determine whether other key genes in cell cycle regulation were disrupted in AAH lesions Kurasono and co-workers (1998) investigated Cyclin D1, *RB* and p16. Cell cycle regulation is reviewed in figure 3. Investigations revealed that cyclin D1 was overexpressed in 47% of low grade AAH, 89% of high grade AAH, 28% CIS and 35% of overt ADC. These results suggested that overexpression of cyclin D1 is a common occurrence in AAH. From these data, one may hypothesise that if AAH lesions are precursor lesions of ADC then *cyclin D1* overexpression is not required for the maintenance of the malignant phenotype. If this was the case then one may expect to see a disruption in a component of the *RB* pathway downstream of cyclin D1 e.g. actually within the *RB* gene or within a cyclin dependent kinase inhibitor such as p16. In this study analysis of lesions for the RB protein by immunohistochemistry revealed that loss of expression is associated with the progression of CIS to overt ADC and may therefore reflect mutations within the gene. Loss of p16 expression was shown to be infrequent in AAH lesions, CIS and overt ADC, results that support previous findings that loss of p16 is associated with metastatic spread (Okamoto *et al.*, 1995; discussed in section 1.3.2.3). An alternatively theory would be that increased

levels of cyclin D1 may be a result of growth arresting conditions rather than growth promoting conditions, since p53 has been shown to induce cyclin D1 synthesis through p21 (Chen *et al.*, 1995). This theory is however not supported by immunohistochemical studies.

Due to the absence of identifiable preneoplastic lesions in pulmonary ADC the genetic pathway of pulmonary ADC was unclear. The proposal that areas of AAH are preneoplastic lesions of parenchymal ADC has changed this and since the completion of my work, the genetic profile of AAH lesions has been extensively analysed. From these studies, one may propose a genetic pathway for parenchymal ADC. Research has shown that the oncogenic activation of *K-RAS* and overexpression of cyclin D1 and *ERB-B2* are early events in ADC tumourigenesis, occurring in low grade AAH lesions. Inactivation of the *RB* gene is thought to occur at a CIS stage whilst mutations within the *p53* gene and allelic loss at 3p and 9p are late events occurring in overt ADC. Loss of p16 expression is associated with metastatic spread.

Despite the identification of known tumour markers in AAH lesions, one must be cautious and cannot still with certainty say that AAH lesions are preneoplastic lesions. The adenoma-carcinoma sequence seen in colorectal cancer has been extensively analysed and evidence suggests that both the accumulation of mutations and the order are important in tumourigenesis (Fearon and Vogelstein 1990; Kinzler and Vogelstein 1996). Mutations within the *K-RAS* gene have been reported to occur in normal colonic epithelium and in foci of hyperproliferating cells that do not lead to colorectal neoplasia (Pretlow *et al.*, 1993; Jen *et al.*, 1994), subsequently Kinzler and Vogelstein proposed the theory of gatekeeper genes. A gatekeeper gene may be described as a gene responsible for maintaining constant cell number in renewing cell populations and ensuring the cell responds to situations requiring net cell growth. Mutation in the gatekeeper gene leads to an imbalance of cell proliferation over cell death. Prior to mutations in the gatekeeper gene, mutations in other genes do not lead to sustainable growth perpetration. In the case of colorectal cancer, the gatekeeper gene is considered to be the *APC* gene. Therefore, in the proposed AAH-parenchymal ADC sequence one may question whether *K-RAS* is a

gatekeeper gene or if mutation within another oncogene or a suppressor gene is required before the presence of *K-RAS* mutations can drive tumourigenesis. Unlike in colorectal cancer, where inactivation of the *APC* gene occurs in >98% of carcinomas (Nishisho *et al.*, 1991; Groden *et al.*, 1991), *K-RAS* is oncogenically activated in approximately one third of parenchymal ADC, this suggests that *K-RAS* is not the gatekeeper gene.

The question of whether these lesions are neoplastic or reactive proliferation of non-neoplastic cells is still unclear despite mounting genetic and morphological studies. Without longitudinal studies to assess the biological and clinical progression of AAH no conclusion can be made.

5. A transgeneic murine model for the analysis of APC function

5.1 Introduction

The *APC* gene is located on human chromosome 5q21. Germline mutations in the *APC* gene characterise the Mendelian dominant inherited disorder, familial adenomatous polyposis coli (FAP) (Bodmer *et al*, 1987; Leppert *et al.*, 1987; Dunlop *et al.*, 1990; Kinzler *et al.*, 1991b). FAP patients develop numerous adenomas within the large intestine, some of which ultimately progress to carcinoma. FAP patients are at increased risk of cancers of the brain, thyroid, bone and focal proliferative lesions of the connective tissue "desmoid tumours", suggesting that the *APC* gene may have a more general role in neoplasia.

In the majority of familial and sporadic colorectal cancer cases both *APC* alleles are shown to be inactivated either by gene mutation alone or by loss of one allele and mutation in the remaining allele (Nishisho *et al.*, 1991; Groden *et al.*, 1991). Biallelic inactivation of the *APC* gene has also been reported in sporadic gastric (Tamura *et al.*, 1994) and hepatic cancers (Oda *et al.*, 1996; Imai *et al.*, 1997). Furthermore, loss of heterozygosity at 5q21 has been observed in sporadic tumours of the breast, ovary, lung and oesophagus (Boynton *et al.*, 1992; Thompson *et al.*, 1993; Kashiwaba *et al.*, 1994). Therefore, it appears that the inactivation of the *APC* gene may be a critical step in the genesis of several types of tumours. This is supported by the fact that the *APC* gene is expressed in a wide range of epithelial, mesenchymal and neuronal cell types (reviewed in section 2.1).

It is currently unclear what precise role APC plays in carcinogenesis. The gene encodes a large protein of 2843 amino acids in which numerous functional domains have been identified (reviewed in chapter 2). The protein can homodimerise, it contains armadillo repeat sequences that are thought to represent a site of protein-protein interaction, two β-catenin binding domains; the constitutive binding domain (consisting of 3 amino acid repeat motifs) and the kinase-regulated binding domain (consisting of 7 amino acid repeat

motifs). Interspersed between motifs within the kinase-regulated binding domain is SAMP (Ser-Ala-Met-Pro) motifs and evidence suggests that the protein Axin interacts with APC through APC's latter five 20 amino acid motifs and that this interaction requires SAMP motifs. The C-terminal portion of the protein contains binding sites for tubulin and EB1, for which a role has been proposed in maintaining microtubule integrity and spindle alignment during cell division and cell movement. APC has also been shown to interact with the human homologue of the *Drosophila* Disc large tumour suppressor protein, which is associated with maintenance of cellular adhesion and possibly with the regulation of cell division and differentiation.

In summary, the APC protein appears to have multiple functions within the cell which involve the regulation of cell proliferation and differentiation, cell adhesion and signalling, and cell movement.

5.1.1 Current murine models for the analysis of APC function

In order to understand the role of APC in tumourigenesis, several different murine models of FAP have been generated. The first model was generated using a random mutagenesis protocol employing ethylnitrosurea, a point mutagen (Moser *et al.*, 1990). Following the treatment of mouse strain C57BL/6 with the mutagen, one mouse displayed an anaemic phenotype which was traced to bleeding from multiple adenomas (in excess of 50) in the intestine. A murine line was established and named Min (multiple intestinal neoplasms). Subsequent characterisation revealed that this mouse harboured a nonsense mutation at codon 850 in the mouse *APC* homologue (Su *et al.*, 1992). The protein product, which would be translated from an *Apc* gene harbouring a truncating mutation at codon 850, would contain only the homodimerisation domain and the armadillo repeats. All β-catenin binding domains would be lost, as would domains for interactions with EB1, microtubules and DLG. Mice which were heterozygous for the mutation ($Apc^{+/Min}$) had a life span of approximately 5 months, after which time the secondary effects of the tumours, anaemia and intestinal blockage, resulted in death (Moser *et al.*, 1990). Breeding Min mice onto different genetic backgrounds revealed that

a smaller number of tumours developed and therefore mice lived longer (Dove *et al.*, 1994; Bilger *et al.*, 1996). This suggested the presence of modifier loci. One modifier loci has been identified and this locus, sited at the distal end of mouse chromosome 4, was termed the *MOM-1* locus (modifier of Min) and is discussed in section 2.4. Recent positional cloning experiments have shown that a secretory phospholipase 2 group 2a is a candidate modifier gene situated in the MOM-1 locus (MacPhee *et al.*, 1995; Cormier *et al.*, 1997). The protein product of Pla2g2a functions in the cleavage of fatty acids from lipids and is therefore thought to effect the cellular microenvironment within the intestinal crypt. *MOM-1* sensitive strains are thought to be null for Pla2g2a protein whereas in resistant strains the protein is present (Cormier *et al.*, 1997). C57BL/6 is a sensitive strain for *MOM-1*.

The phenotype displayed by Min mice differs from the human FAP phenotype. In Min mice tumours appear in the small intestine, whereas in human FAP, the majority of mutations are within the large intestine (Bilger *et al.*, 1996). Secondly, a small proportion of female mice exhibit mammary adenomas and ancanthomas (approximately 5%) (Moser *et al.*, 1993; van der Houven *et al.*, 1995). This is not a phenotype of human FAP, and may suggest mutations in genes in the Min mice other than *Apc*. Alternatively, the susceptibility to breast cancer due to germline mutations in the *APC* gene in humans is not significantly increased above the incidence in the general population.

Desmoid tumours and epidermal cysts are detected in Min mice. These are phenotypically associated with human FAP, however other extracolonic manifestations such as osteomas of the skull and mandibles and hypertrophy of retinal pigment epithelium are not seen in this murine model. The reason for this may be related to the position of the germline mutation (codon 850) since there is some evidence to suggest that the particular site of the *APC* mutation correlates with extracolonic manifestations seen in FAP patients (Nagase *et al.*, 1993; Olschwang *et al.*, 1993; Wallis *et al.*, 1994; Bunyan *et al.*, 1995; Caspari *et al.*, 1995; Dobbie *et al.*, 1996). Alternatively, the different phenotypes displayed between the human and murine model of FAP may suggest that genes other than *APC* may be involved in tumourigenesis of some of the extracolonic manifestations.

Analysis of colorectal tumours which developed in the small intestine of Min mice revealed that the remaining wildtype *Apc* allele was lost (Luongo *et al.*, 1994; Levy *et al.*, 1994). Loss of the remaining wildtype allele or inactivation by gene mutation is also reported to occur in the majority of tumours from FAP patients (Nishisho *et al.*, 1991; Groden *et al.*, 1991; Ichii *et al.*, 1992; Nagase *et al.*, 1993).

The study of Min mice has also shown that Apc is important in development since homozygosity for the mutated *Apc* allele ($Apc^{Min/Min}$) leads to embryonic lethality (Moser *et al.*, 1990; Dove *et al.*, 1994; Moser *et al.*, 1995). Analysis of embryos from Min crosses ($Apc^{+/Min}$ X $Apc^{+/Min}$) showed that by 6.5 days post coitum (dpc) abnormalities could be clearly seen in 25% of embryos and genotyping demonstrated that that abnormalities were associated with homozygosity for the mutated *Apc* allele. Close inspection of 6.5 dpc embryos and more mature embryos demonstrated that the development of the primitive ectoderm fails prior to gastrulation and embryos die before day 8 gestation. By mid gestation (approximately 10 dpc) homozygotes consist of a mass of trophoblast giant cells with an additional cluster of much smaller embryonic cells. This phenotype is similar to that seen in both β-catenin and ε-cadherin null embryos (Haegel *et al.*, 1995; Larue *et al.*, 1994).

Murine models of FAP have also been generated using gene targeting methods (Fodde *et al.*, 1994; Oshima *et al.*, 1995). Oshima and co-workers designed a targeting cassette, which was inserted into exon 15 of the murine *Apc* gene by homologous recombination in embryonic stem (ES) cells. Embryonic stem cells and subsequently a murine founder line were generated which expressed the mutant *Apc* gene and encoded a protein product truncated at residue 716. Mice heterozygous for the mutated allele developed multiple adenomas throughout the entire intestinal tract, with the majority occurring within the small intestine. The number of polyps detected within these mice were more numerous (three to five times) than seen in Min mouse model, although both mice are on a C57BL/6 background. The difference may be a result of the fact that in Oshima's model an additional 133 amino acids of the Apc protein is transcribed compared to the Min mouse. No specific binding domain has been identified within this region.

As seen with the Min model, the colorectal tumours that develop have lost the remaining wildtype *Apc* allele and mice homozygous for the mutated allele ($Apc^{716/716}$) die *in utero* before day 8 gestation. In this model, there have been no reports of extracolonic manifestations and this may support the hypothesis that additional genetic events are required for the development of these tumours.

Fodde and workers have also generated a murine model for FAP. The *Apc* 1638N model carries a targeted frameshift mutation which results in the production of a premature stop codon at codon 1638 (Fodde *et al.*, 1994). The protein translated from an *Apc* gene with a truncating mutation at codon 1638 would retain homodimerisation sites, the armadillo repeat sequences as well as both β-catenin binding domains. However, analysis of 1638N mice failed to detect the predicted 182 kDa truncated protein by western blotting technique, therefore suggesting that this protein was unstable. Previous studies of mutations within the *APC* gene, in both mice and humans, have shown that mutations upstream of codon 1400 result in the production of a stable protein. In these cases high numbers of intestinal adenomas have been detected with a low incidence of extracolonic manifestations (Fodde and Khan, 1995). Mutations beyond codon 1600 do not result in a stable protein and are generally associated with a reduced number of intestinal adenomas and severe extracolonic lesions, a phenotype termed attenuated FAP (Scott *et al.*, 1995; Eccles *et al.*, 1996; VanderLuijt *et al.*, 1996; Friedl *et al.*, 1996; Scott *et al.*, 1996). There is no current explanation for these phenotypic differences, but one may suggest that these phenomena may be a result of tissue specific differences relating to APC function or tumour initiation. Further studies are required, especially in light of the recent identification of *APC* homologues (see section 2.5).

The *Apc* 1638N model is characterised by the attenuated FAP phenotype, a low number of adenomas (5-6 adenomas per animal) are detected within the intestine, and unlike the Min or 716 model, the majority exist in the large intestine. Several other extracolonic manifestations, associated with human FAP have been noted, these include osteomas, dental abnormalities, lipomas and retinal pigment epithelium. Additionally, mammary tumours and trichofolliculmas have been observed sporadically (Fodde *et al.*, 1994; Smits

et al., 1998). This murine model is also on a C57BL/6 background.

Loss of the wildtype *Apc* allele has been observed in intestinal tumours although interestingly the frequency at which the wildtype allele is lost in desmoid tumours and cysts is 42% and 53% of cases respectively. The remaining *Apc* allele may be inactivated by a point mutation within the gene or alternatively these results may suggest that the complete inactivation of *Apc* may not be necessary for desmoid tumour or cyst formation. As seen in the Min and 716 models for FAP, homozygosity for the 1638N truncating mutation results in embryonic lethality before day 8 of gestation.

Following the generation and characterisation of the 1638N murine model, Fodde and co-workers then went on to generate the *Apc* 1638T murine model, which was characterised by a stable 182 kDa truncated Apc protein comprising the amino-terminal 1683 amino acids (Smits *et al.*, 1999). The investigators used gene targeting, using the same cassette used to generate the 1638N mice, but with the selectable marker within the cassette which lay in the same transcriptional orientation as *Apc*. Targeted ES cells and subsequently murine founder lines expressed a stable 182 kDa truncated Apc protein. Analysis of mice heterozygous for 1638T did not show an increased tumour predisposition when compared to wildtype animals and furthermore, mice homozygous for 1638T were viable.

The lack of FAP phenotype indicates that deletion of the carboxyl-terminal domains that bind tubulin, DLG and EB-1 does not predispose to tumour development. The protein produced from 1638T harbours the homodimerisation domain, the armadillo repeats and constitutive β-catenin binding domain and three of the seven motifs within the β-catenin kinase regulated domain, which contain one of the three Axin-binding SAMP motifs. Analysis of 1638T homozygous ES cells and embryonic fibroblasts showed that the regulation of β-catenin appeared to be normal whereas in compound heterozygotes 1638T/1638N and cells homozygous for 1638N, β-catenin levels are deregulated. Therefore, it appears that the multiprotein complex of APC/Axin, GSK-3β and β-catenin is necessary for regulation of intracellular β-catenin levels (see section 2.6.7) and this regulatory activity is the main tumour suppressing function of APC.

Mice which were homozygous for the *Apc* 1638T mutation were characterised by developmental abnormalities such as growth retardation, behavioural abnormalities, reduced postnatal viability in the absence of preputial glands and the formation of nipple-associated cysts. This phenotype may be a reflection of the loss of function of the C-terminal domains.

5.1.2 Generation of murine models using Cre/lox technology

Several murine models of FAP have been generated either by random chemical carcinogenesis or conventional gene targeting (see section 5.1.1). Different phenotypes have been noted in each model enabling further understanding of Apc's role in the tumourigenesis of colorectal cancer and extracolonic manifestations associated with the FAP phenotype. Furthermore, these models have revealed that the β-catenin binding domains are crucial to the tumour suppressing functions of Apc and that the Apc protein plays a crucial role in development. However, the analysis of Apc function in these models is limited due to embryonic lethality. To overcome this problem, murine models need to be generated which allow inactivation of the *Apc* gene in a temporally controlled manner. In such a model, *Apc* could be inactivated in the adult mouse, therefore bypassing embryonic lethality. Secondly, the gene could also be inactivated during stages of embryogenesis to further examine the role of Apc in development. Such a model can be generated using the Cre-loxP recombination system of the bacteriophage P1 (Sternberg, 1981).

5.1.2.1 Cre recombinase mediated DNA excision and inversion

Bacteriophage P1 encodes Cre recombinase (cyclization recombination), a 38 kDa enzyme, which is a member of the integrase family of recombinases. Cre is a site specific recombinase which functions within the bacteriophage to cyclize P1 DNA after infection (Segev and Cohen 1981; Hochman *et al.*, 1983) and enhance P1 plasmid stability in bacterial lysogens by resolving dimeric plasmids prior to bacterial division (Austin *et al.*, 1981). Cre recombinase recognises a 34 bp sequence on the P1 genome called loxP

(locus of X-over of P1). The consensus sequence for the loxP sequence, shown in figure 15a, catalyses reciprocal conservative DNA recombination between pairs of loxP sites. Resultant DNA structure is dependent upon the orientation of loxP sites. Direct repeats of loxP dictate excision of intervening sequences whereas inverted repeats specify inversion of the intervening DNA sequence (Abremski *et al.*, 1983). Cleavage and joining of DNA is confined to discrete positions within the core region of the loxP sequence (the spacer region) and proceeds one strand at a time by way of a transient phosphotyrosine DNA-recombinase complex. In the case of DNA excision, one loxP sequence is also removed, in fact this is composed of half of each of the two loxP sequences which are each cleaved within the spacer region. Excision is reversible and the excised fragment can be reinserted, however, excision is much more efficient than reinsertion (reviewed in Torres and Kuhn, 1997). Inversion is also reversible and occurs at a similar frequency in both directions (reviewed in Torres and Kuhn, 1997). Figure 15b depicts Cre recombinase mediated DNA excision and inversion. Unlike many recombinases of the Int family, no host factors are needed for efficient Cre-mediated DNA recombination, therefore the functions of this recombinase can be put to use in the genomic manipulation of eukaryotic cells (Sauer 1987).

The loxP sequence consists of two inverted 13 bp repeat sequences that flank an 8 bp-spacer sequence (see figure 15a). Not all 34 bp are essential for efficient recombination. The first 4 bp of either of the 13 bp inverted repeats (as read from each end of the sequence) can be modified without significant loss of recombination proficiency or fidelity, each 13 bp sequence binds a single Cre monomer. The asymmetrical 8 bp-spacer region is the site of cleavage and strand exchange of recombining loxP sites (Hoess and Abremski 1985). Experiments have shown that base substitutions can be made within the spacer sequences (Hoess *et al.*, 1986). The following points have been made. Firstly, for efficient recombination to occur between two loxP sites, the spacer region of each loxP sequence must be homologous. Secondly, a base change at position +1 of the spacer region drastically reduces recombination, even between two homologous loxP sequences. Thirdly, if base substitutions are such that the spacer region becomes symmetrical then

these sites undergo both excision and inversion at an equal frequency.

LoxP sequences with a substituted base within the spacer region are of particular use if one is designing an experiment where other genes are also activated and/or inactivated by the use of Cre recombinase. For example, a model may be made whereby two separate genes are "floxed" (insertion of LoxP sites to flank the DNA sequence of interest) and are to be excised following the administration of Cre recombinase. If genes are on the same chromosomal strand then several possible excisions or recombination events may occur, due to the presence of 4 loxP sites. If one gene harbours loxP sites with a base substitution within the spacer region then recombination/excision will occur between this pair and not with either of the loxP sequences containing the consensus sequence. Cre recombinase also catalyses intermolecular recombination resulting in integration and translocation. Therefore, even if the floxed genes are not on the same DNA strand care must be taken in design of the experiment.

Figure 15: The Cre/loxP recombination system

Figure 15a: The consensus sequence for loxP sites.

Cre recombinase recognises a 34 bp sequence called loxP (locus of X-over of P1) and catalyses reciprocal conservative DNA recombination between pairs of loxP sites. The sequence consists of two inverted repeat sequences that flank an 8 bp spacer sequence. Cre recombinase binds to the DNA at the inverted repeat sequences; each repeat sequence binds one Cre monomer. The spacer sequence determines the orientation of the loxP sequence and is the site of enzymatic cleavage and strand exchange of recombining loxP sites.

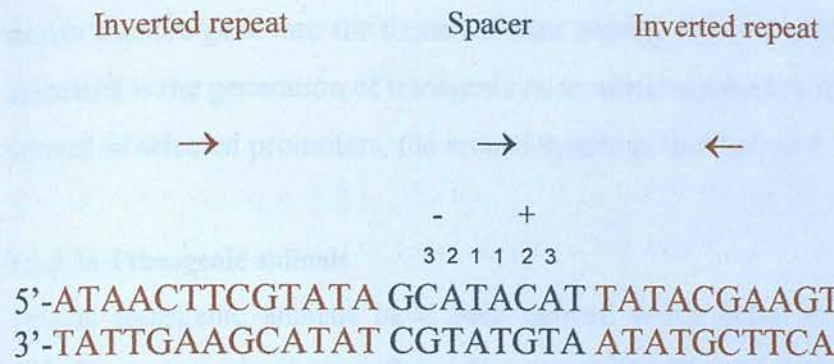
Base changes within the first 4 bp of either of the inverted repeat sequences (as read from the end of each sequence) or within positions –1, -2 –3 or +2,+3 of the spacer region still result in high recombination proficiency and fidelity. In the case of a degenerate sequence within the spacer region, efficient recombination only occurs if this sequence is homologous in both loxP sites.

Figure 15b: Cre/loxP recombination reactions

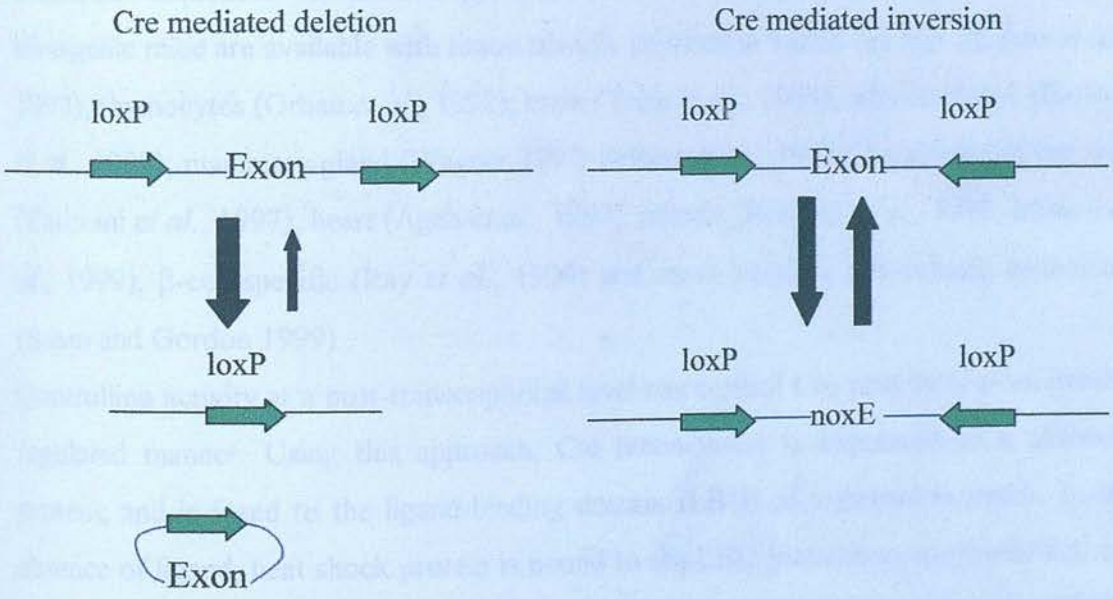
The resultant DNA structures following Cre mediated recombination, is dependent upon the orientation of loxP sites (as determined by the spacer region, see above). Direct repeats of loxP dictate excision and recirculation of intervening sequences, leaving one loxP site on each product. Excision is reversible with the excision reaction being more efficient than reinsertion. Inverted repeats specify inversion of the intervening DNA sequence, this reaction is reversible and occurs at an equal frequency in both directions.

Figure 15: The Cre/loxP recombination system

15a: Consensus sequence for loxP sites



15b: Cre/loxP recombination reactions



5.1.2.2 Delivery of Cre recombinase

Following the generation of a mouse line which harbours the floxed gene, gene expression can be controlled by Cre recombinase. Two systems are currently in use to deliver the Cre gene into the tissues of mice bearing floxed alleles. The first system to be discussed is the generation of transgenic mice which expresses Cre recombinase under the control of selected promoters, the second system is viral delivery.

5.1.2.2a Transgenic animals

Several transgenic animals have been derived which either express Cre recombinase globally or in a tissue specific and/or temporal manner. Cre transgenic mice can be crossed with the floxed transgenic mice, and the gene of interest controlled in doubly transgenic progeny.

For global expression of Cre, the gene is driven by promoter sequences from ubiquitously expressed genes or from viral promoters (Schwenk *et al.*, 1995; Lakso *et al.*, 1996; Saki and Miyazaki 1997). Transgenic mice have also been generated which express Cre recombinase in a tissue specific manner. These models have utilised tissue specific promoter sequences to drive expression of the Cre-recombinase gene. Currently, transgenic mice are available with tissue specific expression within the eye (Lakso *et al.*, 1992); thymocytes (Orban *et al.*, 1992); brain (Tsien *et al.*, 1996); adipose tissue (Barlow *et al.*, 1997); mammary gland (Wagner 1997; Selbert *et al.*, 1998); basal cells of the skin (Tarutani *et al.*, 1997), heart (Agah *et al.*, 1997), muscle (Bruning *et al.*, 1998; Miniou *et al.*, 1999), β-cell-specific (Ray *et al.*, 1999) and small intestine and colonic epithelium (Saam and Gordon 1999).

Controlling activity at a post-transcriptional level can control Cre activity in a temporally regulated manner. Using this approach, Cre recombinase is expressed as a chimeric protein, and is fused to the ligand-binding domain (LBD) of a steroid receptor. In the absence of ligand, heat shock protein is bound to the LBD preventing enzymatic activity of the Cre protein. Upon administration of ligand, heat shock proteins dissociate and Cre

activity is restored. To further control Cre activity and eliminate activity of the enzyme in the presence of endogenous steroid, the LBD may be mutated so that it has a very low affinity for the natural ligand but can bind to synthetic steroids. This system has been successfully used in ES cells and transgenic mice with the LBD of oestrogen and progesterone receptors (Kellendonk *et al.*, 1996; Zhang *et al.*, 1996; Feil *et al.*, 1996). Schwenk and co-workers (1998) have generated a transgenic model whereby Cre activity is both tissue specific and temporally-regulated. In this model, a B-cell specific promoter is used to limit the expression of a Cre-oestrogen receptor fusion protein to B-lymphocytes. This model exhibits both temporal and tissue specific control of Cre recombinase.

5.1.2.2b Viral delivery

An alternative approach for the delivery of the Cre gene into tissues involves the use of replication deficient recombinant viruses. Viral delivery of the Cre recombinase gene into selected tissues is a much more rapid approach than crossing transgenic lines (as described above). A number of Cre-expressing adenovirus vectors have been generated (Anton and Gramham 1995; Sakai *et al.*, 1995; Kanegae *et al.*, 1995; Wang *et al.*, 1996b; Rohlmann *et al.*, 1996). Adenovirus is able to infect a range of cell types making it a powerful vehicle for gene delivery (Brody and Crystal 1994; Smith *et al.*, 1995). Adenovirus may be administered by intravenous injection, studies have shown this route of administration results on high levels of recombination within liver and spleen, with lower levels in the kidney, heart and lung (Wang *et al.*, 1996b). High levels of infection in specific tissues are thought to be a result of some cell types being particularly susceptible to adenovirus infection. In these studies only trace levels of adenovirus were detected in the brain, presumably because the brain-blood barrier blocks viral particles (Doran *et al.*, 1995; Wang *et al.*, 1996b). The expression of Cre may be controlled by placing the gene under a tissue specific promoter, as discussed above, or alternatively the adenovirus may be injected directly into the organ of interest (Wang *et al.*, 1996b; Agah *et al.*, 1997; Stec *et al.*, 1999).

This method of Cre delivery does have some problems, adenovirus can be antigenic and cause cellular immune responses against the viral particles and viral infected cells (Kass-Eisler *et al.*, 1994).

5.2 Aims

In the following sections, I describe the generation and initial characterisation of a new *in vivo* model for the analysis of APC function, which utilises a transgenic Cre/Lox approach. This novel model will allow the analysis of :

1) *APC* transgenic mice on a murine wildtype *Apc* background. These mice will ubiquitously express the *APC* transgene, and transcription of APC will be controlled by the murine phosphoglycerate kinase gene promoter. Essentially the *APC* tumour suppressor gene will be overexpressed throughout the mouse.
2) the ability of the transgene to rescue the embryonic lethal phenotype of mice homozygous for the *Min* mutation.
3) the temporal and spatial deletion of the *APC* transgene.

The *APC* transgene will be engineered so that expression can be conditionally inactivated using the Cre-lox P recombination system of the bacteriophage P1 (Sternberg *et al*, 1981). The transgene constructed for this study requires the incorporation of loxP sites upstream and downstream of the *APC* cDNA sequence, which has previously been cloned into a plasmid. Two loxP sequences will be inserted in the same relative orientation into the plasmid backbone to allow excision of the *APC* cDNA following administration of Cre recombinase. LoxP sequences will harbour a base substitution within the spacer region.

The expression of the transgene will be controlled by the promoter sequence of the murine phosphoglycerate kinase (*Pgk*) gene. The *Pgk* gene is a housekeeping gene and encodes a key enzyme required for glycolysis. This gene is ubiquitously expressed, and particularly high levels are noted in early embryonic cells that are rapidly proliferating, and much lower levels are seen in many differentiated cells (Riele *et al.*, 1990; Mortensen *et al.*, 1991; McBurney *et al.*, 1991; 1994; Sutherland *et al.*, 1995). The murine *Pgk* promoter sequence will be cloned upstream of the *APC* cDNA but downstream of the

first loxP sequence.

Following the construction of the transgene, a prokaryotic system will be used to confirm that the *APC* cDNA can be excised from the construct in the presence of Cre recombinase. Pro-nuclear injection will be used to deliver the *APC* transgene into oocytes of wildtype mice. Mice produced as a result of pronuclear injection will be screened for the presence of the transgene and these founder mice will be employed in a breeding program to produce lines of mice that bear the *APC* transgene on an *Apc* $^{Min/Min}$ background.

5.3 MATERIALS AND METHODS

Construction of the floxed *APC* transgene involved numerous techniques. Each technique will be addressed individually.

5.3.1 Ethidium bromide plates

Ethidium bromide plates were used to determine the concentration of DNA. Preparation of plates and method is as previously described in section 3.3.3.

5.3.2 Agarose gel electrophoresis

All agarose gels were prepared and analysed according to the protocol previously described in section 3.3.4.2

5.3.3 Restriction endonuclease cleavage

Plasmid DNA was cleaved with the appropriate restriction endonuclease required for the cloning step. Reaction mixes were prepared as described in section 3.3.4.3. Concentrations of DNA to be cleaved varied and subsequently the amount of enzyme and volume of the reaction was increased proportionately.

5.3.4 Polymerase chain reaction

Reaction mixes for PCR were prepared as previously described in section 3.3.4.1

5.3.5 Extraction of DNA fragments from agarose gels

Following PCR or restriction endonuclease cleavage, it was sometimes necessary to isolate a particular fragment of DNA from other DNA within the reaction. The required fragments were isolated from the agarose gel and then purified using the QiaexII gel extraction kit (Qiagen Ltd.). The appropriate fragments were cut from the agarose gel using a sterile scalpel blade and placed into a pre-weighed 1.5 ml microfuge tube and

microfuged at 13,000 rpm for 2 minutes. The microfuge tube was then re-weighed to determine the weight of the gel slice. Three hundred microlitres of QX1 buffer was added per 100 mg of gel slice to solubilise the agarose, dissociate DNA binding proteins from the DNA fragment and give a suitable pH to allow absorption of the DNA fragment to the QiaexII silica particles. Ten microlitres of Qiaex II particles were added and the sample incubated at 50°C for 15 minutes. Particles were washed with QX 1 buffer and PE buffer. The pelleted particles were air-dried and the DNA fragment eluted with 20µl of sterile DDW.

5.3.6 Dephosphorylation of vectors

The 5' terminal nucleotide of linearised vectors were dephosphorylated using shrimp alkaline phosphatase (Amersham, UK). One unit of shrimp alkaline phosphatase was added to the 100 ng of linearised vector in reaction buffer (20 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$). The reaction was incubated at 37°C for 1 hour then heated to 65°C to inactivate the enzyme. This step was included to increase ligation efficiency by preventing re-circularisation of the vector.

5.3.7 Ligation reactions

Concentrations of both vector and insert were estimated using ethidium bromide plates (see section 3.3.3). Ligations were prepared using 100 ng of vector and various molar ratios of insert generally 1:1 or 1:3, vector to insert. Conversion of molar ratio to mass ratio was calculated using the following equation:

$$\frac{\text{100 ng of vector x kb size of insert}}{\text{kb size of vector}} \text{ x molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

To each ligation, 1µl of 10x ligation buffer (Appendix A) and 1µl T4 DNA ligase (4.0 Weiss units) (Gibco BRL) was added. The volume of the ligation mix was adjusted to

10µl by the addition of sterile DDW. Ligation reactions were incubated overnight at 16°C. Ligation controls were also set up, these reactions contained no insert but only dephosphorylated vector.

5.3.8 Transformations

Transformations were carried out using Epicurian Coli SURE competent cells (Stratagene Ltd., Cambridge, U.K.). The protocol was according to Stratagene instructions. Competent cells were removed from –70°C and thawed on ice. Fifty microlitres of cells were placed in a pre-chilled 15 ml Falcon tubes and β-mercaptoethanol (supplied with the cells) added to give a final concentration of 25 mM. Cells were incubated for 10 minutes on ice then 1-2µl (5-8µg of DNA) of ligation mix added and the tubes incubated on ice for a further 30 minutes. Tubes were placed in a 42°C water bath for 45 seconds and then immediately placed on ice for 2 minutes. Nine hundred microlitres of preheated (42°C) SOC medium (Appendix A) was added and the tubes incubated at 37°C for 1 hour shaking in an orbital shaker at 225 rpm. One hundred microlitres and 200µl of each transformation was spread onto separate LB (Luria-Bertani) agar plates containing 50µg/ml of ampicillin (Appendix A). Ampicillin selects for bacterial colonies that have successfully been transformed with the plasmid as the plasmid contains the ampicillin resistance gene. The transformation solution was left to absorb into the agar plate for approximately 5 minutes and then plates were inverted and placed at 37°C for at least 18 hours. Plates were then placed at 4°C.

5.3.9 Bacterial cultures and plasmid isolation

To screen for colonies which had been transformed with the plasmid and contained the appropriate insert, individual colonies (clones) from the agar plates were transferred into 3 mls of LB broth supplemented with 50µg/ml ampicillin (Appendix A) and incubated shaking at 225 rpm at 37°C overnight. Plasmid DNA was purified from 2.75 mls of the

bacterial culture using the Wizard™ plus Minipreps DNA Purification System (Promega UK). This protocol is outlined in section 4.3.5.3. The concentration of the plasmid DNA was estimated using ethidium bromide plates (section 3.3.3).

5.3.9.1 Bacterial streak plates and glycerol stocks

Bacterial streak plates and glycerol stocks were prepared from the remaining 250µl of the 3 ml cultures. Bacterial plates were prepared for short-term storage of clones. A biological loop was inoculated with the culture then streaked across LB (Luria-Bertani) agar plates containing 50µg/ml of ampicillin (Appendix A). The plates were incubated overnight at 37°C and then stored at 4°C. For long term storage of clones, 0.1 ml of the culture was pipetted into an eppendorf and 0.1 ml of sterile glycerol added. The cell culture and glycerol were gently mixed and the eppendorf was frozen at -70°C.

5.3.10 Large scale preparation of plasmid

To obtain a large volume of the plasmid for further restriction analysis, cloning steps or plasmid sequencing, a single clone was removed from the streaked agar plate (see section 5.3.9.1) and used to inoculate 3 mls of LB broth supplemented with 50µg/ml ampicillin (Appendix A). The bacterial culture was incubated shaking at 225 rpm at 37°C overnight. This culture was then used to inoculate 3-100 ml of culture medium, the volume of culture used was 1/100th the volume of new culture medium. The inoculated medium was then incubated shaking at 225 rpm at 37°C overnight and Wizard™ midiprep columns were used to isolate plasmid DNA. The protocol is as described in section 4.3.5.3 but with solution volumes scaled up.

5.3.11 Sequencing plasmid DNA

The procedure used to sequence plasmid DNA is described in section 4.3.5.4. Samples were loaded onto a DNA sequencing gel and electrophoresed using conditions described in section 4.3.4.2. Sequencing gels were subsequently dried and exposed to x-autoradiograph film as described in section 4.3.4.2.

5.4 Construction of the *APC* transgene

The construction of the *APC* transgene is described in the following sections and each step is depicted in figure 16.

5.4.1 *APC* cDNA

The human *APC* cDNA sequence was originally generated by screening cDNA libraries with genomic DNA segments from a contig of yeast artificial chromosome (*et al.*, 1991b). The *APC* cDNA sequence is 8532 bp in length, and contains a methionine codon (ATG) at the 5' end of the sequence. This initiation codon is situated two codons from an in-frame stop codon conforming to the Kozak translation initiation consensus (Kozak, 1984). Multiple translation stop sequences are present in all frames over the final 1.8 kb within the 3' open reading frame. Since the isolation of the *APC* cDNA sequence, the sequence has been cloned into the plasmid pGEM-3ZF(+) (p-GEM) by S. White (unpublished). The *APC* cDNA, cloned into the *BamH1* restriction site within the multiple cloning site of pGEM-3ZF(+), was available in our department for this project, as shown in figure 16a. This construct formed the backbone to the *APC* transgene.

5.4.2 Addition of the SV40 poly-adenylation sequence

A poly-adenylation sequence was not present at the 3' terminus of the *APC* cDNA sequence. This nucleotide sequence results in the transcription of 100-200 residues of adenylic acid (poly-A tail) to the 3' end of the RNA chain in order to complete the RNA transcript. The function of the poly-A tail is not known for certain but is thought to play a role in the export of mRNA from the nucleus and also to stabilise mRNA molecules by retarding their degradation in the cytoplasm.

The first step was to clone in a poly-adenylation sequence. PCR was used to amplify the SV40 poly-adenylation (poly-A) sequence from the plasmid pUHD 10-3 (constructed by M. Gossen, Heidelberg, unpublished). Degenerate primers were designed which

incorporated *Mlu1* sites into the amplified fragment to generate compatible ends for the cloning of the poly-A sequence into a unique *Mlu1* site situated 3' to the *APC* cDNA stop codon, see figure 16b.

Primers were designed by S. White and resulted in the production of a 480 bp PCR fragment. Primers are stated below.

Forward 5'- GGCG***ACGCGT***AGACATGATAAG 3'

Reverse 5' - ATCG***ACGCGT***AAGCTTGGTCGA 3'

The underlined nucleotides indicate SV40 sequence, italics indicates *Mlu1* restriction recognition sequence. PCR conditions for these primers were: initial denaturation of 5 minutes at 94°C, and then 30 cycles of 94°C for 1 minute, 60°C for 30 seconds and 72°C for 1 minute, the final extension time was 72°C for 8 minutes. Reaction mixes for PCR were as previously described in section 3.3.4.1 with the exception of the use of Expand™ high fidelity PCR system (Boehringer Manheim). Expand high fidelity PCR system was used in order to minimise the possibility of polymerase errors within the PCR product. The Expand high fidelity PCR system utilises thermostable *Taq* DNA and *Pwo* DNA polymerases (Barnes, 1994). Due to the 3'-5' exonuclease proofreading activity of *Pow* DNA polymerase, the fidelity of the DNA synthesis is increased 3 fold compared to *Taq* DNA polymerase alone.

Two *Mlu1* restriction endonuclease reactions were prepared, one containing the PCR amplified poly-A sequence, the other the *APC* cDNA construct (as seen in figure 16a). Following restriction endonuclease cleavage, ligation reactions were prepared (see section 4.3.5.1) and subsequently transformed into Epicurian Coli SURE competent cells (Stratagene Ltd., Cambridge, U.K.) as described in section 5.3.8. Plasmid DNA was isolated from bacterial clones, sees section 4.3.5.3, and screened for the presence of the poly-A sequence. Screening was carried out by incubating the plasmid preparations with the restriction endonuclease *BstE1*. The restriction endonuclease site is unique to the poly-A sequence. Therefore, following incubation with the restriction endonuclease, if the poly-A sequence is present, one would expect the plasmid to linearise to give a 12 kb fragment whilst in the absence of the sequence the plasmids remains circularised.

Restriction endonuclease reactions were loaded onto a 0.7% agarose gel, a molecular weight marker was also loaded. Marker XII (Boehringer Manheim) was used as this contained a DNA fragment of 12 kb. The samples were electrophoresed until the marker DNA had been separated. Plasmid preparations which linearised were identified.

To determine the orientation of the poly-A sequence, linearised plasmid (previously incubated with *BstN1*) were incubated with the restriction endonuclease *Sal1*. A *Sal1* restriction endonuclease site is present within the p-GEM plasmid backbone. If the poly-A sequence was inserted in the correct orientation, the *Sal1* cleavage site would be approximately 380 bp downstream of the *BstN1* cleavage site. However, if the poly-A sequence has been cloned into the plasmid in the incorrect orientation the *Sal1* site would be situated 80 bp downstream of the *BstN1* restriction site. Therefore, following incubation with *Sal1* samples were electrophoresed through a 3% agarose gel. All plasmid preparations showing a fragment of 380 bp were selected.

One hundred millilitre cultures were grown and plasmid isolated for each selected clone as described in section 5.3.10. Plasmid DNA was sequenced (see section 4.3.5.4) to confirm the presence of the poly-A sequence and its orientation. Sequencing was initiated from the universal M13 reverse sequence situated approximately 80 bp downstream of the poly-A sequence within the p-GEM plasmid backbone. The universal M13 reverse primer sequence was as follows: 5' CAGGAAACAGCTATGAC 3'.

5.4.3 Addition of LoxP sequence

Oligonucleotides were designed which contained a 34 bp loxP sequence with a base pair change at position 2 of the spacer region (discussed in section 5.1.2.1) and compatible ends for sub-cloning into the *APC* cDNA constructs as shown in figure 16b. The first loxP (Lox P 1) to be inserted into the construct was to be cloned into a unique *Sal1* restriction endonuclease site situated within the p-GEM plasmid backbone downstream of the poly-A sequence, see figure 16c. The oligonucleotide Lox P 1 was designed to include a unique restriction site, *Pac1*, which would enable screening of clones and also

Sal1 cohesive ends to allow easy cloning of the oligonucleotide into the *Sal1* restriction site of the *APC* cDNA construct. Oligonucleotide sequences are shown in table 10. The Lox P 1 oligonucleotides were made doubled stranded by mixing an equal molar ratio of each oligonucleotide and heating to 75°C, mixes were left to cool at RT for approximately 15 minutes.

The *APC* cDNA construct, as seen in figure 16b, was incubated with the restriction endonuclease *Sal1*. Ligation reactions were prepared using the cleaved *APC* cDNA construct and the annealed Lox P1 oligonucleotides. Ligation mixes were subsequently transformed, bacterial cultures grown and plasmid DNA isolated using the previously described standard methodology. Plasmid DNA was subjected to restriction endonuclease cleavage using the enzyme *Pacl*. Restriction endonuclease reactions were then electrophoresed through a 0.7% agarose gel. Plasmid preparations viewed as a 12.2 kb linear fragment were sequenced to ensure the oligonucleotide had not integrated as a concatamer and to determine orientation. Sequencing was carried out using the M13 reverse primer previously described in section 5.4.2.

A second Lox P oligonucleotide was designed to be cloned into the *Sac1* restriction endonuclease site present within the p-GEM plasmid backbone upstream of the *APC* cDNA sequence, see figure 16d. This oligonucleotide (Lox P 2) was designed with *Sac1* cohesive ends and an *Apa1* restriction site downstream of the 34 bp loxP consensus sequence, see table 10. The *APC* cDNA construct, as seen in figure 16c, was incubated with the restriction endonuclease *Sac1*. Ligation reactions were prepared using the cleaved *APC* cDNA construct and annealed Lox P 2 oligonucleotides (see above for oligonucleotide annealing conditions). Ligation, transformations and DNA plasmid preparations were carried out as described for Lox P 1. Isolated plasmid DNA was incubated with the restriction endonuclease *Sac1* and electrophoresed through a 0.7% agarose gel. Plasmid preparations that linearised, giving a 12.2 kb fragment, harboured the Lox P 2 sequence. These plasmid preparations were sequenced to ensure the oligonucleotide had not integrated as a concatamer and to determine orientation. Sequencing was carried out using the M13 universal forward primer (5'

GTTTTCCCAGTCACGAC 3') which was situated 74 bp upstream of the inserted Lox P 2 sequence.

Table 10. Oligonucleotide sequences for Lox P 1 and Lox P 2

Oligonucleotide sequences and orientation within plasmid	
Lox P 1: 3' to the *APC* cDNA ← Inverted Repeat Spacer Inverted Repeat 5' *G*TATTGAAGCATAT *TACATATG* ATATGCTTCAATA**AATTAATT***CAGCT* 3' 3' *TCGAC*ATAACTTCGTATA *ATGTATAC* TATACGAAGTTAT**TTAATTAA***G* 5'	***Sal1 cohesive ends*** **Pac1 restriction site**
Lox P 2: 5' to the *APC* cDNA → Inverted Repeat Spacer Inverted Repeat 5' **C**CGGCCGATAACTTCGTATA*ATGTATAC*TATACGAAGTTAT*GGGCCC***GAGCT** 3' 3'**TCGAG**GCCGGCTATTGAAGCATAT*TACATATG*ATATGCTTCAATA*CCCGGG***C** 5'	**Sac1 cohesive ends** ***Apa1 restriction site***

5.4.4 Addition of the murine phosphoglycerate kinase gene promoter sequence

The final step in constructing the *APC* transgene was the addition of the phosphoglycerate kinase (Pgk) gene promoter sequence. The murine Pgk promoter sequence was available in our department within a 4 kb construct, Puc19-Pgk, previously prepared by S. White. The 900 bp Pgk promoter sequence could be excised from the Puc19-Pgk plasmid by incubation with the restriction endonucleases *BamH1* and *Pst1*. The Pgk sequence was to be cloned into the unique *Apa1* restriction endonuclease site within the *APC* construct as shown in figure 16d. Therefore, to aid cloning, two sub-cloning steps were carried out. One microgram of the Puc19-Pgk plasmid was linearised using the restriction endonuclease *BamH1* and ligated with *ApaI/BamHI* linkers. *Apa1/BamH1* linkers are oligonucleotides consisting of the endonuclease recognition sequences for *BamH1* and *Apa1*. The oligonucleotide sequence consisted of *BamHI/Apa1/BamHI* nucleotide sequences. Oligonucleotide sequences were made double stranded as described in section 5.4.3, and the ligation reaction between the linearised Puc19-Pgk plasmid (incubated with *BamH1*) and oligonucleotides carried out. Following the ligation reaction, the mixes were transformed, colonies selected, bacterial cultures prepared and plasmid DNA isolated, using standard methodology. Each plasmid preparation was screened for the presence of the linker by incubation with *Apa1*. Plasmid DNA from clones containing the linker linearised with *ApaI*.

Plasmid DNA, from a clone containing the *Apa1* linker, was then cleaved with the restriction endonuclease *Pst1* and an *Apa1/Pst1* linker, (nucleotide sequence *Pst1/Apa1/Pst1*) cloned into the plasmid as previously described above. Plasmid DNA from clones were screened by *Apa1* cleavage, excising the 900 bp Pgk promoter sequence with *Apa1* cohesive ends. For clones shown to harbour the Pgk promoter, large-scale plasmid preparations were prepared (10 ml), as described in section 5.3.10. The 900 bp Pgk promoter sequence was isolated from the PUC19 plasmid backbone (size 3 kb) by incubating 1μg of the isolated plasmid with *Apa1* and electrophoresis through a 0.7% agarose gel and subsequent gel extraction using the method descried in section 5.3.5.

The isolated Pgk sequence containing *Apa1* cohesive ends was sub-cloned into the *APC* construct at the *Apa1* site, using standard techniques. To select for clones that contained the Pgk promoter sequence, plasmid preparations were screened by incubating with *Apa1*. Positive clones were detected by gel electrophoresis (0.7% agarose gel), a fragment of approximately 900 bp corresponded to the presence of the Pgk promoter sequence and a 12 kb fragment the *APC* construct as seen in figure 16d. Plasmid DNA from clones containing the Pgk promoter sequence were screened by *Pst1* cleavage to determine the orientation of the promoter sequence. The Vector NTI software package was used to determine the predicted pattern of plasmid fragments that would be generated following incubation with *Pst1*. The fragment pattern expected from *Pst1* cleavage is shown below.

Pgk promoter correct Orientation	Pgk promoter incorrect orientation
4673	4673
4117	**3217**
2163	2163
2164	1361
-----	**900**
480	480
400	400

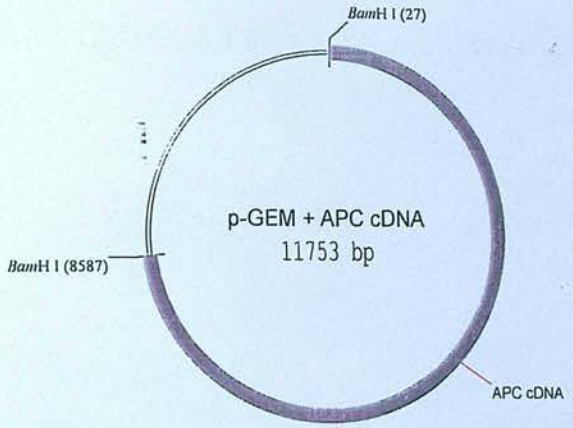
Plasmid preparations were subsequently incubated with *Pst1* and reaction mixes electrophoresed through a 0.7% agarose gel. One plasmid preparation containing the Pgk promoter sequence in the correct orientation was identified, as seen in figure 16f, and this construct is the completed *APC* transgene. To produce large amounts of the transgene, four 100 ml bacterial cultures were set up and plasmid isolated according to the method in section 5.3.10. Multiple glycerol stocks were also prepared as described in section 5.3.9.1.

Figure 16: Construction of the *APC* transgene

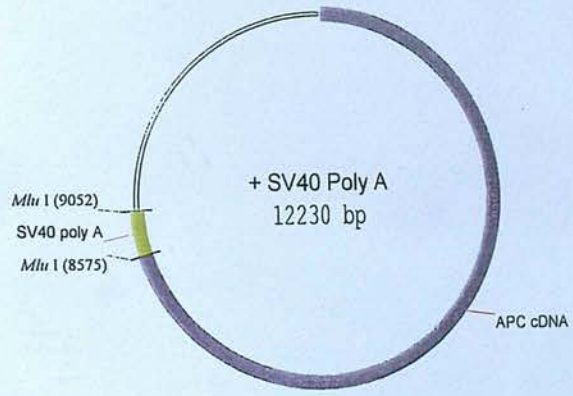
Each cloning step is described within the text of section 5.4.1-5.4.4.

Figure 16: Construction of the *APC* transgene

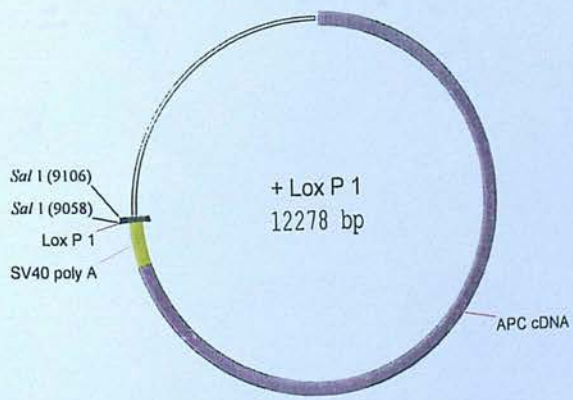
a)



b)



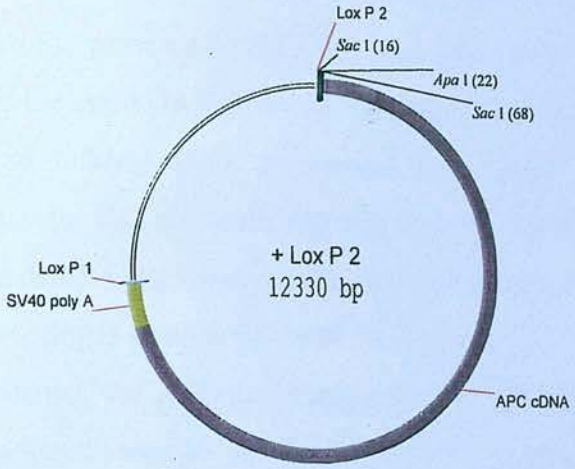
c)



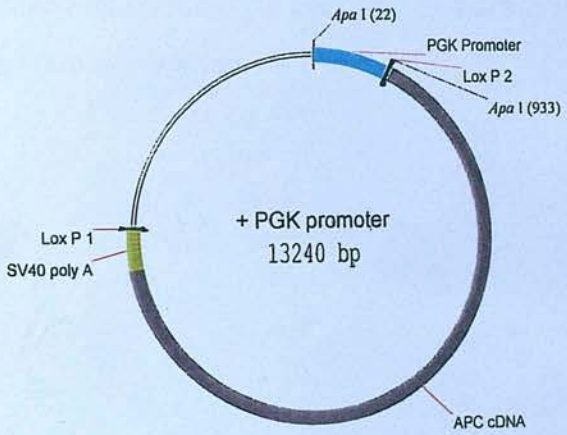
Continued overleaf

Figure 16: Continued from overleaf

d)



e)



f)

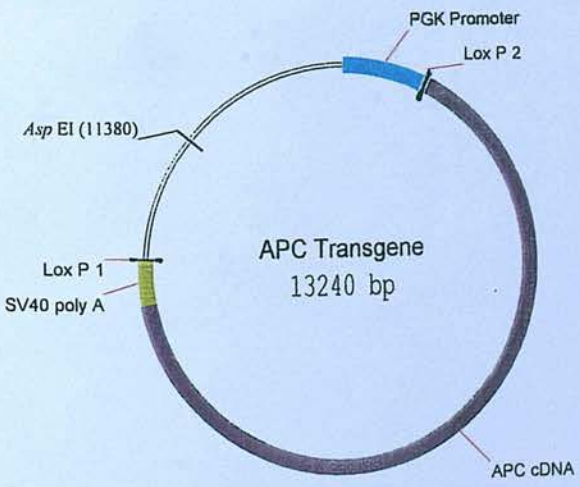


Figure 16 continued: Construction of the *APC* transgene

Each cloning step is described within the text of section 5.4.1-5.4.4.

5.5 Analysis of the *APC* transgene

The construction of the *APC* transgene is described in sections 5.4.1-5.4.4. Restriction endonuclease reactions and/or DNA sequencing was carried out following each cloning step to ensure the presence of the DNA sequence being inserted and to determine the orientation of the insert. In the case of the addition of oligonucleotide sequences (Lox P sequences and linkers) DNA sequencing was carried out not only to determine orientation but to also eliminate the presence of oligonucleotide concatamers which would not be detected by fragment sizes following agarose electrophoresis. An example of DNA sequencing is given in figure 17.

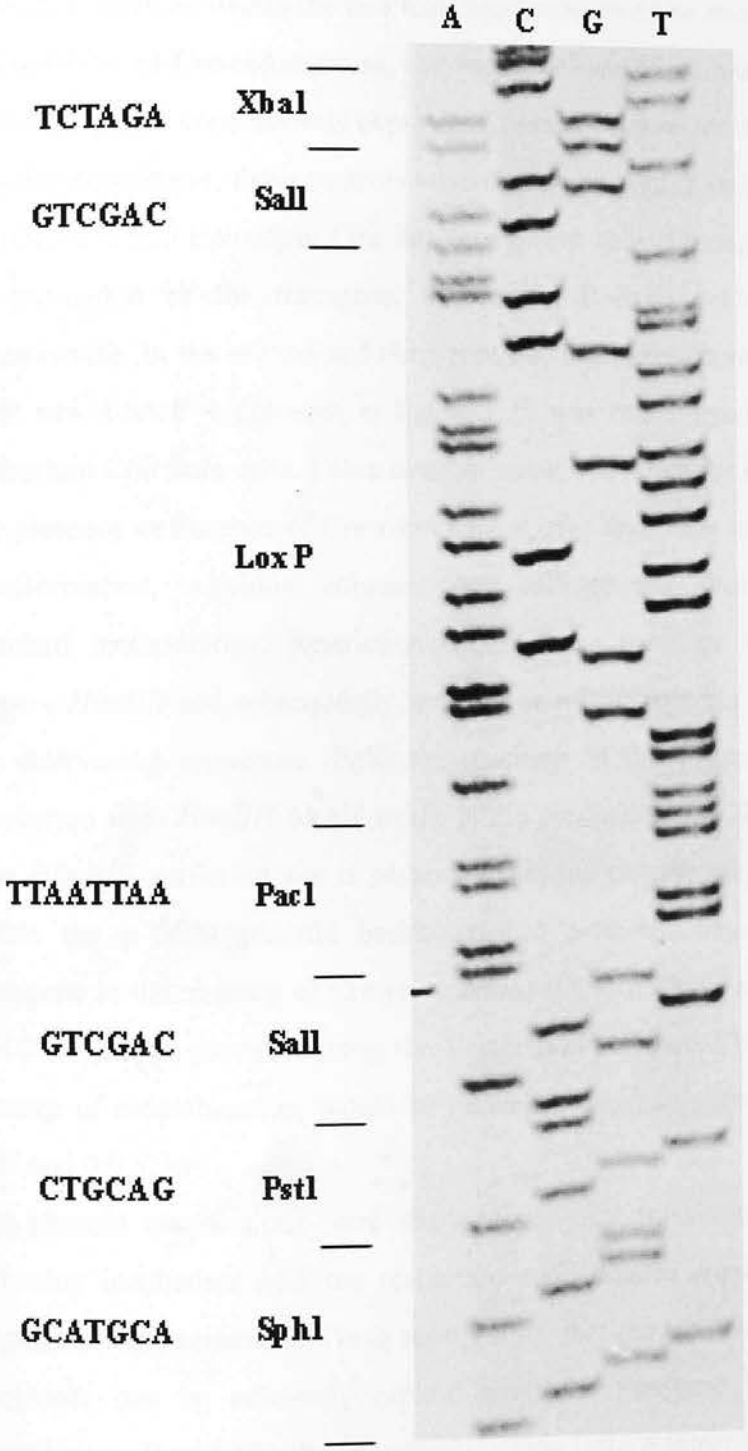
The final construct, the *APC* transgene, was then incubated with a number of different restriction endonucleases to confirm the relative positions of the cloned fragments. Digestion patterns were analysed on a 0.7% agarose gel and compared to the expected pattern determined by building the construct on the Vector NTI software package. All fragment sizes were as predicted.

Figure 17: Example of DNA sequencing

The Lox P 1 sequence was cloned into the *Sal1* restriction endonuclease site within the *APC* construct as seen in figure 16c. Clones containing the sequence were identified by restriction endonuclease cleavage with *Pac1*. Plasmid DNA was subsequently sequenced to determine the orientation of the Lox P 2 sequence and to ensure that concatamers had not been formed.

Prior to insertion of the Lox P 2 sequence the DNA sequence encoded *Sph1*, *Pst1*, *Sal1* and *Xba1* restriction endonuclease recognition sequences. Following the cloning of the Lox P1 sequence into the *Sal1* restriction site, *Sph1*, *Pst1* and *Xba1* surround one copy of the Lox P 1 sequence.

Figure 17: Example of DNA sequencing



5.5.1 Excision of *APC* by Cre-recombinase in prokaryotic cells

To test whether the DNA sequence (*Pgk* promoter sequence, *APC* cDNA and poly-A sequence) situated within the two loxP sequences could be excised from the transgene in the presence of Cre-recombinase, the transgene was transformed into the E. *Coli* strain BNN132, which constitutively expresses Cre recombinase (MoBiTech, GmbH).

For this experiment, three controls were also carried out. Firstly, the *APC* transgene was transformed into Epicurium Coli Sure competent cells (Stratagene), to demonstrate that recombination of the transgene, as seen in BNN132 cells, was a result of Cre recombinase. In the second and third controls, incomplete construct containing only one loxP site, Lox P 1 (as seen in figure 16c) was transformed into BNN132 cells and Epicurium Coli Sure cells. These controls would show that for recombination to occur, in the presence or absence of Cre recombinase, two loxP sites must be present. Following transformation, individual colonies were cultured and plasmid DNA isolated using standard methodology. Restriction endonuclease reactions were prepared using the enzyme *HindIII* and subsequently analysed on a 0.7% agarose gel to identify excision of the intervening sequences. Following excision of the sequence between lox P sites, incubation with *HindIII* would result in the production of a single fragment of 3.2 kb. One *HindIII* restriction site is present within the recombined construct and is situated within the p-GEM plasmid backbone. The predicted fragment sizes for the *APC* transgene in the absence of Cre recombinase (control 1) are 6.382, 3.863, 2.630, 0.276 and 0.096 kb, as predicted using the Vector NTI software. The control construct, in the absence of recombination, would be cleaved to yield fragments of 5.420, 3.863, 2.630 0.27 and 0.096 kb.

Ten plasmid preparations were analysed for each transformation, in all preparations following incubation with the restriction endonuclease *HindIII* the appropriate sized fragments were generated. These results show that the DNA sequence between the loxP sequences can be efficiently excised from the transgene in the presence of Cre-recombinase. Figure 18a shows examples of the DNA fragments identified.

Figure 18a: Cre mediated excision of the floxed *APC* construct in a prokaryotic system

Figure 18a shows that the DNA sequence between the loxP sequences can be efficiently excised from the *APC* transgene in the presence of Cre recombinase.
Lane 1 shows the DNA fragments generated following *HindIII* cleavage of plasmid preparations generated from the transformation of the *APC* transgene into Epicurium coli Sure cells, this strain does not express Cre recombinase. Fragments of approximately 6.3, 3.8 and 2.6 kb are seen. Further fragments of 276 and 96 bp are generated but are not detected on this agarose gel.
Lane 2 shows the DNA fragments generated following *HindIII* cleavage of plasmid preparations generated from the transformation of the *APC* transgene into an E. coli strain (BNN132) which constitutively expresses Cre recombinase. Following Cre mediated excision, a 3.199 kb construct is generated, and this construct contains a single *HindIII* site, following incubation the construct is linearised.
Lane 3 and 4 show the DNA fragments generated when a control construct, containing only one loxP site, is transformed into Epicurium sure cells (lane 3) and BNN132 cells (lane 4). These lanes show that in the absence or presence of Cre recombinase this construct does not undergo recombination. In both lanes, as predicted, DNA fragments of 5.4, 3.8 and 2.6 kb are seen.

Figure 18b: Cre mediated excision of the floxed *APC* construct in an eukaryotic system

Transgene positive embryonic fibroblasts were derived and infected with a replication deficient adenovirus that either expressed Cre recombinase (flask 1), or β-galactosidase (flask 2). Flask 4 was not infected with adenovirus.
DNA was extracted from each flask and subject to PCR to amplify a 240 bp fragment of the transgene. Primers were situated within the poly A sequence and the p-GEM plasmid backbone. Following Cre mediated excision of the floxed sequence, the poly A sequence would not be present and no PCR product would be produced. Results are shown in figure 18b (i). A 240 bp PCR product was detected in lanes 1,2 and 3 (flask 1, flask 2 and flask 4 respectively). No band was seen in lane 4, the PCR with no DNA template control. The presence of a 240 bp PCR product in lane 1 indicated that Cre mediated recombination may not have occurred or was not 100% efficient. Alternatively, reinsertion of the excised fragment had occurred.
DNA was then subject to PCR to determine whether the transgene had under gone Cre mediated excision. The primers used spanned the floxed sequence, if excision had occurred, a 190 bp fragment would be amplified. Results are shown in figure 18b (ii). A 190 bp fragment was detected in lane one; DNA extracted from flask 1; infected with adenovirus expressing Cre recombinase. No band was detected in lanes 2,3 and 4, infected with adenovirus expressing β-galactosidase, not infected and PCR no template control.

Figure 18a: Cre mediated excision of the floxed *APC* construct in a prokaryotic system

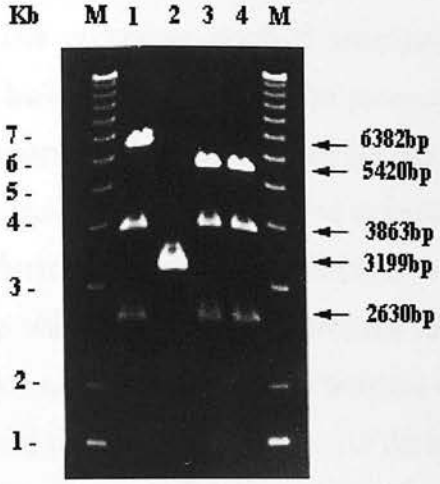
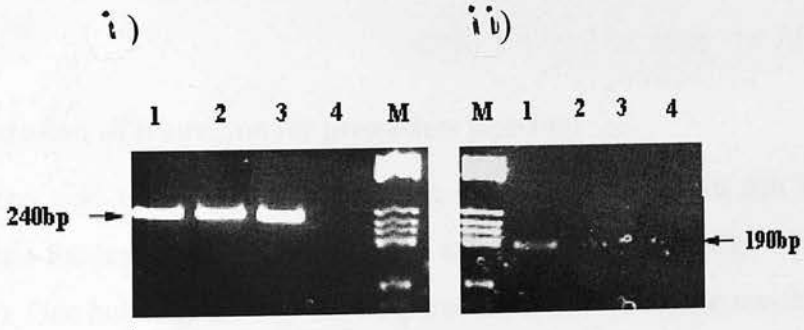


Figure 18b: Cre mediated excision of the floxed *APC* construct in an eukaryotic system



5.6 Generation and analysis of transgenic mice

The following sections describe the generation and analysis of a murine *APC* transgenic founder line. The procedure involved linearisation and purification of the transgene, which was to be microinjected into the pronuclei of fertilised oocytes at the one-cell stage. The transgene will randomly integrate into the murine genome, and may integrate as multiple tandemly repeated copies and at multiple sites. Injected oocytes were placed within the oviducts of surrogate female mice who carried the developing pups. Following birth, each pup was analysed for the presence of the transgene using PCR to amplify a region of the transgene; DNA extracted from tail biopsies was used as template. Southern blotting was used to confirm the presence of the transgene and to determine the transgene copy number. Expression of the transgene was analysed by reverse transcription PCR. Transgene positive mice were bred to establish *APC* transgenic founder lines. Transgene positive mice essentially overexpress *APC*, therefore, tissues were analysed to detect any phenotype associated with aberrant expression.

Transgenic founder lines were used in a breeding program with the aim of generating transgene positive mice on a *Apc* $^{Min/Min}$ background, i.e. the transgene will rescue of the embryonic lethal phenotype of *Apc* $^{Min/Min}$ mice.

5.6.1 Preparation of transgene for pronuclear injection

The first step was to linearise the transgene as it has been shown that linear DNA integrates into the host genome fivefold more efficiently than supercoiled DNA (Townes *et al.*, 1985). One hundred and eight micrograms of the *APC* transgene was linearised in a total volume of 1ml using the restriction endonuclease *AspE1*. The nucleotide sequence recognised by *AspE1* endonuclease is situated at position 11380 of the *APC* transgene, see figure 16f. Following overnight incubation at 37°C, heating the reaction mix to 65°C for 15 minutes inactivated the restriction endonuclease. The *APC* transgene was purified by phenol chloroform extraction and ethanol precipitation. The first step was to add an equal volume (1ml) of phenol: chloroform: isoamyl alcohol (25:24:1) solution to the

sample. The solution was mixed well, then microfuged at 13,000 rpm for 1 minute. The aqueous phase (upper layer) was pipetted into a clean microfuge tube whilst the cloudy interface, containing the proteins and the lower organic phase, was discarded. An equal volume of chloroform was then added and the sample spun at 13,000 rpm for 1 minute. The upper phase, containing the transgene, was pipetted into a clean microfuge tube and the organic lower phase discarded. The purified transgene was then concentrated and sterilised by ethanol precipitation. For ethanol precipitation, 2 volumes of ethanol and $1/10^{th}$ the volume of 3M sodium acetate pH 5 were added to the DNA solution. The sample was placed at -20°C for at least 1 hour and then centrifuged at 13,000 rpm in a microfuge for 20 minutes. Ethanol was poured off and the pellet washed with 1 ml of 70% (v/v) ethanol. The sample was centrifuged for 15 minutes and the wash step repeated. The transgene was resuspended in sterile injection buffer (10 mM Tris, 0.1 mM EDTA) at a concentration of 6μg/ml. The resuspended transgene was then dialysed against a large volume of injection buffer for 4 days at 4°C; the buffer was changed daily. This step was carried out to remove all contaminating solutes and solvents that might be deleterious to the oocyte. Following dialysis, the concentration of the transgene was adjusted to 10 μg/ml and was stored at –20°C in 20μl aliquots.

5.6.2 Collection of fertilised oocytes and pronuclear injection

Dr. Alistair Mackenzie carried out this work in the MRC Transgenic Facility, Edinburgh. The first step was to treat a 4-5 week old female mice, strain C57BL/6, with follicle stimulating hormone (FSH) (Sigma, UK) and human chorionic gonadotrophin (hCG) (Sigma, UK) to induce superovulation. Mice were injected intraperitoneally (ip) with 100μl of 50 IU/ml FSH (ie 5 IU/mouse). Between 46-48 hours later, the mice were injected ip with 100μl of 50 IU/ml hCG (ie 5 IU/mouse). The female mice were then placed with a stud male, murine strain CBA. C57BL/6 strain was not used, as males tend to be less fertile than other strains. The following morning, females were examined for the presence of a copulatory plug and plugged females were killed by cervical dislocation.

The oviducts were removed and placed into a petri dish containing M2 medium (Sigma, UK) which had been equilibrated to RT. The ampulla, which contains the cumulus mass (numerous fertilised eggs surrounded by cumulus cells), was located by examining the oviducts under a dissection microscope. The ampulla was teased out of the oviduct and the cumulus mass and any isolated fertilised eggs transferred using a pipette containing M2 medium into a second dish of M2 medium. The cumulus masses were mixed with 50µl of hyaluronidase solution (10 mg/ml of hyaluronidase in M2 medium) for a period of 2-3 minutes; this step separated the fertilised eggs from cumulus cells. Fertilised eggs were transferred, using a transfer pipette, to a petri dish containing droplets of M2 medium and washed by repeatedly collecting and transferring to a new droplet of medium in the dish, they were then transferred to a dish containing M16 medium (Sigma, UK) which had been equilibrated to 37°C in a $C0_2$ incubator. Fertilised eggs were washed twice in M16 medium, as described above, and subsequently placed at 37°C in 5% CO_2 until required for microinjection (usually within an hour, within this time the transgene is equilibrated to RT and the micro-manipulator prepared).

The transgene, prepared as described in section 5.6.1, was equilibrated to RT. Fertilised eggs were transferred into M2 medium which had been equilibrated to RT and the transgene was injected into a pronuclei of the fertilised egg using a micro-manipulator. For an in depth description of the use of a micro-manipulator see Hogan *et al* (1994). Briefly, the transgene DNA is forced into one of the two pronuclei of the fertilised egg at the one-cell stage via an injection needle linked to a simple system consisting of an air filled tube connected to a 50 ml syringe. Usually the male pronuclei are targeted, as they tend to be larger than female pronuclei. Enough transgene was injected to see the pronuclei swell but not burst.

Injected eggs were then transferred into equilibrated M16 medium and placed at 37°C in a 5% $C0_2$ incubator. The following day embryos which were at the two cell stage were selected and transferred into the oviducts of 0.5 day post-coitum pseudo-pregnant mice. These mice were prepared by hormonal treatment as described for superovulation, but mature females were used that had already successfully raised a litter of pups. Following

hormone treatment, females were placed with vasectomised males, these males leave a copulatory plug after mating but fail to transmit sperm. Pseudo-pregnant recipient females were anaesthetised and embryos transferred into the ampullae by the oviduct transfer procedure as described in Hogan *et al* (1994). Between 10-18 embryos were transferred to each oviduct in an attempt to generate litter sizes of 5-10 pups. Usually less than a third of transferred embryos result in the production of a live pup. Following this procedure, the mice were returned to a cage and left undisturbed to recover from the anaesthetic. In this experiment, 110 two cell stage embryos were transferred into four recipients.

Recipient mice gave birth to a total of 13 pups, unfortunately two litters (six pups) were lost due to water run outs within the cages. The remaining 7 pups were subsequently screened for the presence of the transgene.

5.6.3 Screening for transgenic mice

5.6.3.1 Extraction of DNA from mouse tails

Tail biopsies were taken from mice of 3 weeks of age and DNA extracted using the following method. Tails were incubated shaking overnight in 500µl of lysis buffer (Appendix A) with 30µl of proteinase K (stock solution of 20 mg/ml). The DNA was purified by phenol extraction. Five hundred microlitres of TE saturated phenol was added and the tube shaken vigorously. The sample was then microfuged for 3 minutes and the aqueous phase carefully removed to a clean microfuge tube, avoiding the lower phenol phase and material at the interface. This procedure was followed again for 500µl of phenol: chloroform (1:1) and then 500µl of chloroform. DNA was precipitated by the addition of 200µl of 7.5M ammonium acetate and 3.75 ml of absolute alcohol. The samples were shaken and a glass rod was used to spool out the DNA. Excess alcohol was allowed to evaporate from the DNA before it was transferred to 500µl of TE buffer (Appendix A). DNA samples were stored at 4°C.

5.6.3.2 PCR protocol

PCR was carried out on the extracted DNA samples to identify mice that contained the *APC* transgene. To control against false negatives, all DNA extracts were initially subjected to PCR to amplify a 144 bp fragment of the murine *Apc* gene. Primers and PCR conditions have previously been published (Luongo *et al.*, 1994).

To detect the presence of the transgene, primers were designed that would amplify a 240 bp fragment by PCR. The forward primer was situated in the transgene poly-A sequence and the downstream primer was the universal M13 reverse primer situated within the p-GEM plasmid part of the transgene (see section 5.4.2).

Primers were: forward 5' CTCTTCATCGGGAATGC 3'

reverse 5' CAGGAACAGCTATGAC 3'

PCR cycling conditions were 94°C for 5 minutes for 1 cycle, 34 cycles of 94°C for 1 minute, 55°C for 30 seconds and 72°C for 30 seconds. Final extension time was 8 minutes at 72°C.

5.6.3.3 Results

Of the 7 pups screened, two were positive for the *APC* transgene, see figure 19. Unfortunately, one transgene positive pup was found dead (pup 7). This pup died aged 4 weeks and had been noted to phenotypically resemble a runt. It was not possible to histologically analyse this pup, as when it was found the tissues were substantially decayed. Thus, there was only one founder mouse to establish one transgenic line.

Figure 19: Detection of *APC* founder lines

The progeny of surrogate mothers, who under went oviduct transfer for oocytes injected with the *APC* transgene, were screened for the presence of the transgene. DNA was extracted from tail biopsies of the 7 pups and subject to PCR to amplify a 240 bp fragment of the transgene. PCR reactions were electrophoresed through a 3% agarose gel containing ethidium bromide. Lanes 1 to 7 represent PCR reactions for progeny 1 to 7. Lane 8 is a no DNA template control for the PCR reaction and lane M contains DNA molecular weight marker V. Two of the seven progeny were shown to be transgene positive (lanes 2 and 7).

Figure 20: Detection of RNA transcribed from the *APC* transgene

RNA was extracted from numerous tissues of transgene positive mice. RNA was subject to reverse transcription (rt) to generate cDNA, and samples treated with DNase to remove any contaminating DNA. To determine whether rt had been successful, a fragment of the β-actin gene was amplified by PCR; presence of cDNA resulted in a 460 bp PCR product. PCR reaction mixes were electrophoresed through a 3% agarose gel containing ethidium bromide. Figure 20a shows that in all tissues analysed, a PCR product of 460 bp was present. Lanes were as follows:

M-represents molecular weight marker V

1: kidney	6: small intestine	11: brain
2: spleen	7: lung	12: thymus
3: pancreas	8: heart	13: mammary gland
4: muscle	9: testis	14: salivary gland
5: large intestine	10: liver	

15: positive control cDNA

16: negative control; cDNA derived from the tail of a non-transgenic animal

To detect transcription from the *APC* transgene, cDNA was subject to PCR to amplify a 137 bp fragment encompassing the 3' terminus of the *APC* gene and the transcribed region of the poly A sequence. To control for amplification of contaminating genomic DNA, which would also yield a PCR product of 137 bp, for each tissue a control sample was also analysed. In these samples the cDNA had not been subjected to rt. Reverse transcribed and non rt samples were loaded adjacent to each other on a 3% agarose gel. A 137 bp PCR product was detected for each tissue sample that was subject to rt whilst no bands were detected in the corresponding control samples, as seen in figure 20b and c. Lanes 1-14 are loaded as described for figure 20a. M represents molecular weight marker V.

Figure 19: Detection of *APC* founder lines

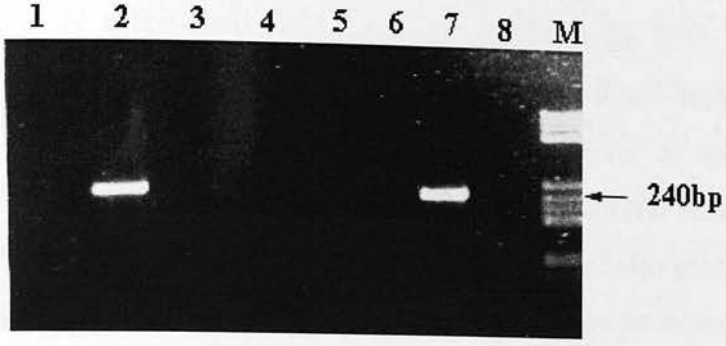
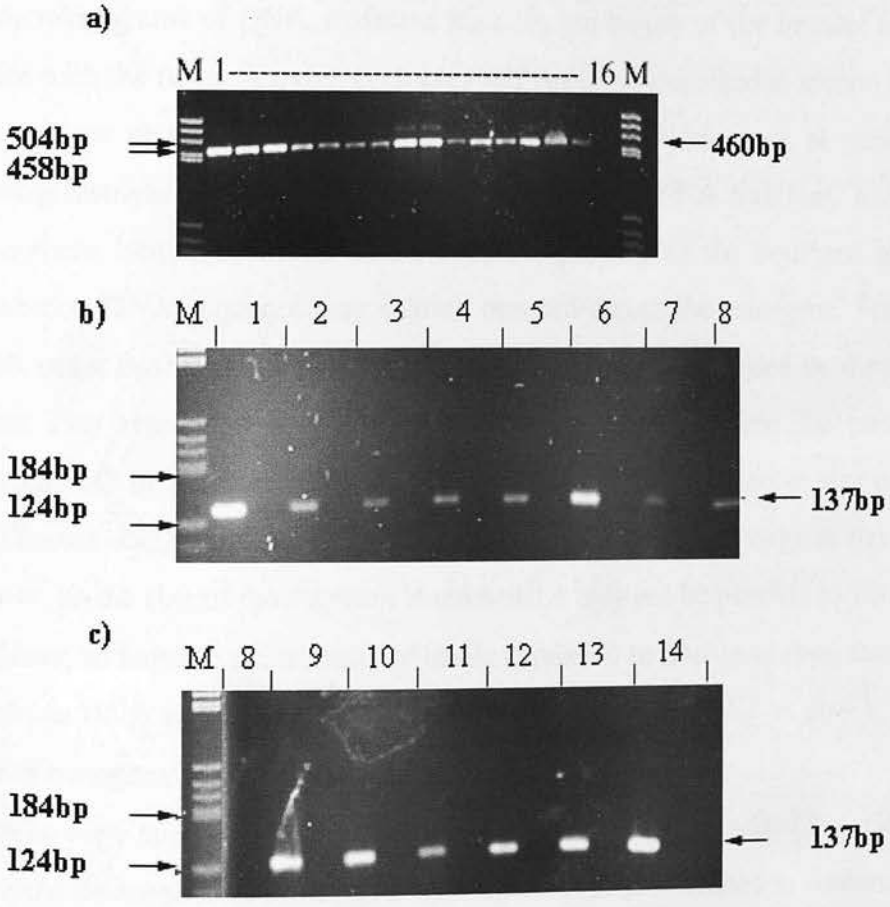


Figure 20: Detection of RNA transcribed from the *APC* transgene



5.6.4 Determination of transgene copy number by Southern blotting

Prior to microinjection into fertilised oocytes, the transgene was linearised (see section 5.6.1). Injection of a linearised DNA constructs results in the DNA being integrated randomly into the genome, the DNA may integrate as a single copy or as multiple tandemly repeated copies, furthermore integration may occur at multiple sites. To determine the transgene copy number and whether integration had occurred at multiple sites the Southern blot protocol was used. This method would also confirm the presence of the transgene to support the PCR result that may give rise to false positives despite negative controls being incorporated.

5.6.4.1 DNA digestion and Southern transfer

Twenty micrograms of DNA, extracted from the tail biopsy of the founder mouse, was digested with the restriction endonuclease *Pml1* (methods described in section 5.6.3.1). A single enzyme recognition sequence is present within the transgene at position 3165. Following restriction endonuclease cleavage of the murine DNA with *Pml1* and following the Southern blotting protocol and subsequent probing of the Southern blot with a radiolabelled DNA sequence (see below), one can detect the transgene. Fragments of 13.2 kb and a further fragment of unknown size, which is determined by the distance to the first *Pml* restriction site in the murine genome from where the transgene has integrated, will be detected using the probe. The fragment of unknown size can be used as an internal control indicating, by the intensity of the band, one copy of the transgene. However, as the size of the fragment is unknown it may not be possible to detect. In this experiment, to indicate the intensity of bands in relation to transgene copy number, 20μg of wildtype DNA was spiked with transgene linearised with *AspE1* to give 1, 10 or 100 copies of transgene per copy of mouse genome.

Transgene copy number control samples and the 20μg of the founder DNA cleaved with *Pml1* were electrophoresed through a 0.5% agarose gel as described in section 5.3.2 with the exception that electrophoresis was carried out overnight at 25 volts. DNA molecular weight marker XII was also electrophoresed through the gel to indicate the position of

the 13.2 kb fragment. The agarose gel was then placed on an UV transilluminator to ensure that all the DNA, and particularly the high molecular weight DNA, had migrated into the gel and was significantly separated. A photograph was taken of the gel with a ruler placed beside the molecular weight marker lane so that following blotting and probing the Southern blot the size of bands detected could be determined.

The following day DNA within the gel was transferred onto a Nylon hybridisation membrane following a modified protocol previously described by Southern (1975). The gel was washed in 0.2M HCl for 10 minutes. This step was incorporated to depurinate the DNA to enable better transfer of the high molecular weight DNA. The gel was rinsed in DDW and then denatured by washing in 0.4M NaOH for 40 minutes, rinsed and washed for 2x 20 minutes in 1M Tris pH 7.2. DNA was transferred from the gel to a nylon membrane by capillary action using the following apparatus. A tray was filled with 20 times SSC (Appendix A) and a perspex plate suspended over the reservoir. A piece of Whatman 3MM paper was cut to the width of the gel and laid over the perspex plate. The length of the paper was such that it acted as a wick dipping into the SSC solution. The gel was placed reverse side up on top of the wick and a piece of Nylon membrane, the same size of the gel pre-soaked in 2x SSC, placed on top. Two pieces of Whatman 3MM paper were then placed on top of the membrane followed by about 2 inches of paper towels. A 1 kg weight was placed on top and the blotting apparatus left overnight. The following day the membrane was marked for orientation and washed in 2 X SSC and baked for 15 minutes at 100°C. The membrane was then stored at RT until it was probed.

5.6.4.2 Labelling DNA probes

The PCR amplified Poly-A sequence (see section 5.4.2) was used to probe the membrane. The PCR product was purified though a wizard™ clean up column to remove any unincorporated nucleotides. One microlitre of Wizard™ clean up resin was pipetted into a 1.5 ml microcentrifuge tube and 5µg of PCR product added. The resin and PCR product were mixed by gently inverting the tube several times. The solution was then pipetted into a syringe barrel, which was attached to the Wizard minicolumn. The syringe

plunger was inserted into the barrel slowly, pushing the solution into the column. Two millilitres of 80% (v/v) isopropanol was added and pushed through the column using the syringe plunger. The syringe barrel was then removed from the minicolumn, and the column seated in a 1.5 ml microcentrifuge tube. The minicolumn was centrifuged at 13,000 rpm for 2 minutes and then transferred to a new microfuge tube. Fifty microlitres of prewarmed (65-70°C) TE buffer (Appendix A) was added to the column. Following a period of approximately 1 minute, the minicolumn was centrifuged for 20 seconds at 13,000 rpm to elute the bound DNA fragment. The concentration of the PCR product was determined and adjusted to a concentration of 200 ng/µl. The purified PCR product was radiolabelled using High Prime kit (Boehringer Mannheim). High prime is a random primer DNA labelling method originally developed by Feinberg and Vogelstein (1983) whereby a complementary DNA strand is synthesised by Klenow polymerase using the 3'OH termini of random oligonucleotides as primers and a labelled deoxyribonucleoside-triphosphate.

Seven hundred and fifty nanograms of purified PCR product was denatured by boiling for 15 minutes and chilling on ice. The volume was adjusted to 11µl and 4µl of high prime mix (Appendix A) and 5µl of ^{32}P dCTP added. The reaction mix was incubated at 37°C for 10 minutes. The labelling reaction was stopped by the addition of 2µl of 0.2M EDTA (pH 8.0).

5.6.4.3 Hybridisation of Southern membranes

Hybridisation was carried out in a Hybaid hybridisation oven with rotisserie. Membranes were placed in the Hybaid bottles and incubated for 2 hours at 65°C in 15 ml of pre-hybridisation solution (Appendix A) containing 100µg/ml of denatured sonicated salmon sperm DNA. The probe was then denatured by boiling for 5 minutes and placed on ice prior to being added to the pre-hybridisation solution to hybridise overnight. Following incubation, the hybridisation solution was discarded and the membranes washed at 65°C with the following different stringency washes; 2 x 15 minutes with wash 1, 1x 30 minutes with wash 2 and 1 x 10 minutes with wash 3. For autoradiography the membrane

was wrapped in Saran Wrap and exposed to Kodak Biomax MR film at -70°C for a period of between 1-4 weeks. Autoradioactive film was subsequently developed using an automated film developer.

5.6.4.4 Southern blot results

Using the protocol described in preceeding sections I could not determine the transgene copy number or confirm the presence of the transgene. No bands were seen in the lane loaded with the founder line DNA and no bands were seen in the transgene control copy numbers 1 and 10. A band of approximately 13 kb was seen in the lane containing the transgene control copy number 100 lane. This indicated that the sensitivity of this protocol was not capable of detecting less than 100 copies of transgene and that the founder line probably contained less than 100 copies. Increasing amounts of DNA were subject to restriction endonuclease cleavage and loaded onto gels (40μg and 60μg). The only band to be detected again was a 13 kb band in the transgene control copy number 100 sample.

Southern blotting and probing using a ^{32}P radionucleotide probe of between 400-6000 bp has been reported to easily detect single-copy genes in 10μg of human DNA and in practice detection in 0.5μg of total DNA is possible (Old and Primrose 1989). It is currently unclear why this protocol was unsuccessful and this issue will be discussed further in section 5.7.

5.6.5 Establishing the *APC* transgene line

The *APC* transgene positive mouse (the founder mouse) was used to establish the *APC* transgenic line. Initially, the founder mouse, a female, was mated with a male mouse heterozygous for the Min mutation ($Apc^{+/Min}$; see section 5.1.1). Eleven progeny were produced which consisted of four transgene positive (Tg +) mice, three of which were homozygous for wildtype *Apc* (ie Tg +, $Apc^{+/+}$), one which was heterozygous for the Min allele (ie Tg+, $Apc^{+/Min}$). F1 Tg+ mice were then utilised in a breeding program to

generate Tg+, *Apc* $^{+/Min}$ mice, which were subsequently interbred in an attempt to generate Tg+, *Apc* $^{Min/Min}$ mice. The result of this breeding program are presented and discussed in section 5.6.9.

5.6.6 Analysis of transgene expression in transgene positive mice

Whilst the breeding program was underway, experiments were carried out to determine whether RNA was being transcribed from the transgene. Three F1 mice with the genotype 3 Tg+; *Apc*$^{+/+}$ were culled at the age of 2 months. Mice were dissected and tissue taken for RNA extraction. Tissue was placed immediately in liquid nitrogen. A sample of each tissue was also taken for histological analysis and was placed directly in formalin prior to paraffin embedding.

5.6.6.1 RNA extraction from mouse tissues

To prevent degradation of RNA, in the following protocol all solutions used are RNAase free. Double distilled water has been treated with DEPC (see appendix A) and where possible tissues and extracted samples are maintained at 4°C. All equipment that came into contact with the tissue or extract was new or had been treated with DEPC water.

Each tissue for RNA extraction was placed, still frozen, into 1 ml of TRIzol Reagent (Life technologies). TRIzol reagent is a mono-phasic solution of phenol and guanidine isothiocyanate, which maintains the integrity of the RNA while disrupting cells and dissolving cell components. Tissues were homogenised and incubated for 5 minutes at RT. Two hundred microlitres of chloroform was added and the tubes shaken vigorously for 15 seconds and incubated at RT for 3 minutes. The samples were centrifuged at 13,000 rpm for 15 minutes at 4°C, and the aqueous phase containing the RNA was removed to a clean tube. Five hundred microlitres of isopropanol was added to each tube and the samples incubated at RT for 10 minutes. Precipitated RNA was recovered by centrifugation at 13,000 rpm for 15 minutes at 4°C. The RNA pellet was washed in 1 ml of 75% ethanol (prepared with DEPC treated water), centrifuged at 6,500 rpm for 5 minutes at 4°C and the supernatant discarded. The pellet was air dried for 10 minutes and

resuspended in 500µl of DEPC treated DDW and incubated for 10 minutes at 55°C. The yield and purity of the extracted RNA was determined using a spectrophotometer. The yield was determined at 260nm where 1 absorbance unit (A_{260}) = 40µg /ml of RNA and purity was estimated by the absorbance ratio at 260nm/280nm. An absorbance ratio between 1.7-2.0 was considered acceptable. To ensure there was no contaminating DNA, each sample was treated with DNase (Gibco, UK). The protocol was as follows. Two micrograms of RNA was incubated for 15 minutes at RT with 2µl of 10X DNase reaction buffer (Appendix A), 2µl (1 unit/µl) of DNase 1 enzyme and DEPC treated DDW to a final volume of 20µl. DNase was inactivated by the addition of 2µl of 25 mM EDTA and incubation at 65°C for 10 minutes. Samples were centrifuged and placed on ice.

5.6.6.2 Reverse transcription for the production of cDNA

One microlitre of oligo (dT) primer (500µg/ml) (Gibco) was added to 10µl of the DNase treated RNA and the sample heated to 70°C for 10 minutes then placed on ice. Four microlitres of 5x first strand buffer (Appendix A), 2µl of 0.1M dithiothreitol and 2µl dNTPs (10 mM each deoxynucleotide) were added and samples heated to 42°C for 2 minutes. Two microlitres of superscript II enzyme (200u/µl) (Gibco) was added and incubated at 42°C for 50 minutes then 70°C for 15 minutes.

5.6.6.3 PCR to detect transcription of the *APC* transgene

Prior to PCR amplification of the *APC* cDNA, to control against false negative results, all samples were subjected to PCR using primers that would amplify a 460 bp fragment of β-actin cDNA. The primers span an intron and if DNA were amplified rather than cDNA then a 1 kb PCR product would be generated. The PCR cycling conditions used for amplification of the cDNA are such that DNA is not amplified.

upstream 5' CTCCGGCATGTGCAAAG 3'

downstream 5' CGTAGATGGGCACAGTG 3'

PCR conditions were 94°C for 5 minutes, followed by 95°C for 45 seconds, 55°C for 1 minute and 72°C for 1 minute for 35 cycles, and a final extension time of 72°C for 10

minutes.

Figure 20 shows the 460 bp fragment amplified from cDNA for the β-actin gene in all tissues analysed. This showed that RNA had been successfully extracted from the tissues and, following treatment with reverse transcriptase, cDNA had been produced that could be amplified in PCR reactions.

cDNA was subjected to PCR to amplify a 137 bp fragment that encompassed the 3' terminus of the *APC* gene and the transcribed region of the poly A sequence. To control for amplification of contaminating genomic DNA, which would also yield a PCR product of 137 bp, RNA prepared following the protocol in section 5.6.6.1, was not subject to reverse transcription.

Primers were upstream 5' CAGCACAGAATCCAGTGGAAC 3'

downstream 5' CACTGCATTCTAGTTGTGGTTTG 3'

PCR conditions were 94°C for 5 minutes, followed by 34 cycles of 94°C, 58°C and 72°C all for 30 seconds, and a final extension time of 72°C for 9 minutes. Figure 20a shows that a fragment of 137 bp was amplified from all tissues analysed and therefore RNA was being transcribed from the transgene. All controls, samples not subject to reverse transcription, were negative indicating that the bands seen are not a result of amplification of any contaminating DNA.

5.6.7 Analysis of transgene positive embryonic fibroblasts

To confirm that the DNA sequence between the loxP sites could be excised from the transgene in an eukaryotic system, transgene positive embryonic fibroblasts (EFs) were derived and infected with a replication deficient adenovirus that carried the gene encoding Cre-recombinase. To indicate that the adenovirus infection had been successful, EFs were also infected with a replicative deficient adenovirus carrying the

β-galactosidase gene. The production of the bacterial enzyme β-galactosidase (β-gal) can be detected using a chromogen substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). X-gal is cleaved by β-galactosidase releasing a blue indolyl derivative. In the following protocols, EFs were cultured and all incubations were carried out at 37°C in a

5% CO_2 humidified incubator unless stated otherwise. All solutions used were sterile.

5.6.7.1 Derivation of Embryonic Fibroblasts

A transgene positive male mouse was crossed with a transgene negative female. The mothers was sacrificed at 13 days post-coitum and the uterus removed and placed into a petri-dish containing PBS. Embryos were isolated from the uterus and the yolk sac and placenta of each embryo removed and discarded. Each embryo was decapitated and placed in 70% ethanol for approximately 2 minutes to reduce the risk of microbial contamination during culturing. Embryos were placed into sterile universals containing PBS and 2.5% trypsin and were then homogenised using dissecting scissors, and incubated for 15 minutes. Following incubation, 8 mls of EF culture medium (Appendix A) was added to each universal to neutralise the trypsin. The embryonic cells were centrifuged for 3 minutes at 1,100 rpm, the supernatant removed and the cells resuspended in 20 ml of EF culture medium and placed into a 75cm^2 tissue culture flask (Costar). The flasks were incubated at 37°C in a 5% CO_2 humidified incubator.

5.6.7.2 Culturing, and freezing embryonic fibroblasts

Embryonic fibroblasts were cultured until confluent and then harvested by trypsinisation. This involved aspirating off the culture medium and adding 5 mls of PBS containing 0.25% trypsin and 0.02% EDTA to each flask. Flasks were incubated at RT for approximately 5 minutes (until the EFs had dissociated from the bottom of the flask) and the cell suspension was then pipetted into sterile universals. EFs were centrifuged for 5 minutes at 1,100 rpm. The supernatant was removed and the cell pellet resuspended in 5 mls of PBS. The number of cells per flask was determined by standard procedure using a haemocytometer. Some of these cells were stored for further experiments whilst DNA was extracted from the remaining cells (see below).

5.6.7.3 Freezing embryonic fibroblasts

For each embryo, 2 vials of EFs were frozen and stored for further experiments. The

following procedure was used; 1 X 10^6 cells were removed, pelleted by centrifugation at 1,100 rpm and resuspended in 2 mls of EF freezing medium (Appendix A). One millilitre of the suspended cells was pipetted into each cryovial (Costar) and placed at –70°C.

5.6.7.4 DNA extraction and PCR protocol to determine the transgene status of embryonic fibroblasts

DNA was isolated from EFs to determine transgene status. 1 x10^6 cells prepared as described in section 5.6.7.1 and 6.6.7.2, were centrifuged at 1,100 rpm and resuspended in 400µl of DNA lysis buffer (see appendix A). DNA extraction was carried out following the protocol in section 3.3.3, without methods for paraffin removal. Transgene status was determined using PCR, primers and conditions are shown in section 5.6.3.2.

5.6.8 Excision of floxed *APC* cDNA in embryonic fibroblasts

5.6.8.1 Adenovirus infection

Embryonic fibroblasts derived from a transgene positive embryo were removed from –70°C. The cryovial was thawed rapidly by placing in warm water. Following thawing EFs were removed from the vial and placed in a sterile universal with 10 ml of EF culture medium (Appendix A). Cells were gently mixed by inverting the universal and then pelleted by centrifugation at 1,100 rpm for 5 minutes. The EFs were resuspended in another 10 ml of media and centrifuged again. This washing procedure was carried out to remove all traces of freezing medium as it contains DMSO which is toxic to the cells. The EFs were then resuspended into 1 ml of EF culture medium and the number of cells per flask determined by the use of a haemocytometer. Four x 10^5 cells were seeded into each of four 25cm^2 tissue culture flasks, 4.75 ml of EF culture medium was added and flasks were incubated overnight to allow the cells to attach to the bottom of the flask. The following day the old EF medium was aspirated off and the following solutions, as stated in table 11, were applied to each flask in a volume of 1 ml of fresh EF medium.

Table 11: Infection of embryonic fibroblasts with adenovirus

Flask 1	1.2×10^7 pfu of Adenovirus/Cre recombinase	Demonstrate excision
Flask 2	1.2×10^7 pfu of Adenovirus/ β-galactosidase	Control for excision
Flask 3	1.2×10^7 pfu of adenovirus/ β-galactosidase	Demonstrate infection
Flask 4	1 ml of EF culture medium	Control for excision

The multiplicity of infection (MOI) required to give a high percentage of adenovirus infection had previously been determined by Elizabeth Lovejoy to be 50 (ie 50 particle forming units (pfu)/cell). Therefore, when infecting 2.4×10^5 cells, 1.2×10^7 pfu of adenovirus was required.

Flasks were incubated for 1 hour with occasional shaking. Following incubation, the EF medium was removed and 5 ml of fresh EF medium added to each flask. Flasks were incubated for 48 hours. After 48 hours, fibroblasts from flasks 1, 2 and 4 were harvested by trypsinisation and DNA extracted following the protocols outlined in section 5.6.7.

Flask 3 was analysed to determine whether the adenovirus infection had been successful. Embryonic fibroblasts were washed twice with 5 mls of PBS. One millilitre of freshly prepared cell fixative (Appendix A) was added to the flask and incubated for 4 minutes. The fixative was removed and the flask was rinsed twice with 5 ml of PBS. Two millilitres of β-galactosidase stain (Appendix A) was then added and the flask incubated overnight. The following day the stain was removed, 5 ml of PBS added and the cells were examined under the microscope. In virtually all of the EFs the blue indolyl derivative, which results from the cleavage of X-gal by β-galactosidase, was noted. The derivative was present in both the cytoplasm and nuclei of EFs. I estimated that 80% of EFs had been infected by adenovirus carrying the β-galactosidase gene.

5.6.8.2 Detection of Cre recombinase mediated excision

DNA was extracted from flasks 1, 2 and 4, as described in section 5.6.7.4. PCR was used to determine that Cre mediated excision had occurred. The DNA template derived from each flask was subjected to PCR to amplify the murine *Apc* gene to ensure that the

extracted DNA was amplifiable. PCR primers and conditions have previously been described in section 5.6.3.2. In all cases, the PCR was successful.

The DNA was then subjected to PCR to detect the presence or absence of the transgene. The forward primer was situated in the transgene poly-A sequence and the downstream primer was the universal M13 reverse primer situated within the p-GEM plasmid part of the transgene. PCR primers and conditions have been previously described, see section 5.6.3.2. If the *APC* transgene is present a 240 bp PCR product is produced, one would expect to see amplification of a 240 bp fragment in the two control flasks, flask 2 (infected with adenovirus expressing β-galactosidase) and flask 4 (not infected with adenovirus). No band was expected to be seen in Flask 1 that was infected with adenovirus expressing Cre recombinase as the sites for the primers would have been excised following Cre mediated recombination. A 240 bp fragment was detected following amplification of DNA derived from all 3 flasks, see figure 18b (i). As the fragment was detected in flask 1 this suggested that Cre mediated recombination had not occurred, or had not occurred in all cells. The latter is likely as I had previously estimated that 80% of embryonic fibroblasts were infected with adenovirus using this protocol (see section 5.6.8.1). Alternatively, this result may implicate reinsertion of the excised fragment (discussed in section 5.1.2.1).

PCR was then used to detect excision of the DNA sequence that lies between the two loxP sites within the transgene. Primers used were the universal M13 forward and M13 reverse primers (see section 5.4.2 and 5.4.3) situated at the 5' and 3' end of the linearised construct (*AspE1*). If the *APC* cDNA were excised then a PCR product of 190 bp would be produced. If however, the sequence was not excised then the distance for the amplification would be greater than 13 kb and would not be detected by the PCR.

The conditions for PCR were 94°C for 5 minutes, followed by 34 cycles of 95°C for 30 seconds, 55°C for 30 seconds, 72°C 30 seconds. A final extension time of 72°C for 8 minutes was included. Following the PCR reaction, samples were electrophoresed through a 3% agarose gel. A 190 bp fragment was produced from the DNA template derived from flask one indicating that Cre mediated recombination had occurred. No

band was seen following amplification of DNA derived from the control flasks, flasks 2 and 4. Results are shown in figure 18b (ii).

This experiment has shown that Cre mediated excision is occurring between the two loxP sites within the *APC* transgene, however, from these experiments I can not determine with what efficiency excision is occurring. As the transgene has probably integrated into the mouse genome in a tandemly repeated array, the transgene may be lying in a tail to tail or head to head orientation. If this was the case then numerous possibilities exist for excision and inversion events, due to the orientation of the loxP sites. This experiment does not address this issue and will be discussed further in section 5.7.

5.6.9 Summary of breeding program

After 11 months, 155 mice had been generated. Thirty two had the genotype Tg+; $Apc^{+/+}$, 58 Tg-; $Apc^{+/+}$, 43 Tg+; $Apc^{+/Min}$ and 22 Tg-; $Apc^{+/Min}$. Raw data are shown in Appendix E.

Of the 32 Tg+; $Apc^{+/+}$, three were culled for detection of transgene RNA (see section 5.6.6) and 1 for the derivation of embryonic fibroblasts (see section 5.6.7). Three mice had died due to the mother not feeding. The age of the remaining 25 mice ranged from 178-305 days. To date these mice, who should overexpress *APC*, appear phenotypically normal.

Forty-three mice had a genotype Tg+; $Apc^{+/Min}$. Four were culled for the production of embryos (see section 5.6.10), and one had died within one month; phenotypically this mouse was a runt. Twenty-two mice had the genotype Tg-; $Apc^{+/Min}$. Genotypes are shown in table 12.

Table 12: Summary of genotypes

Genotype	No. of Mice	Comments	Final No. of mice
Tg+; *Apc* $^{+/+}$	32	Cull 3; detection of transgene RNA. Cull 1; derivation of embryonic fibroblasts. 3 died; mother not feeding.	25
Tg-; *Apc* $^{+/+}$	58		58
Tg+; *Apc* $^{+/Min}$	43	Cull 4; Analysis of embryos. 1 died; runt.	38
Tg-; *Apc* $^{+/Min}$	22		22

Ideally, to determine whether the *APC* transgene could reduce or eliminate the phenotype seen in *Apc* $^{+/Min}$ mice, one could compare survival of Tg+; *Apc* $^{+/Min}$ and Tg-; *Apc* $^{+/Min}$ mice. However, a direct comparison can not be made because the age and the genetic background of these mice vary, being either 75% or 87.5% C57BL/6. Previous studies have shown that the number and age at which the colonic adenomas appear is related to the genetic background (discussed in section 5.1.1). Therefore, the data are presented taking these two factors into account. Raw data are shown in Appendix E, and is summarised in table 13. Results included in this table are until 1st May 1999. Maintenance of the transgenic colony and harvesting of the mice as from the 4th December 1998 were carried out by Owen Sansom and Jason Reed.

Table 13: Analysis of *Apc* heterozygous mice. Figures represent the number of mice that were culled due to illness/total mice in particular litter.

	Background				Comment
Age of litter (days)	75% C57BL/6		87.5% C57BL/6		
	Tg+; $Apc^{+/Min}$	Tg-; $Apc^{+/Min}$	Tg+; $Apc^{+/Min}$	Tg-; $Apc^{+/Min}$	
177	0/2				
178	0/3				
186	0/2				
189	0/5				
198	0/1	0/1			
202	0/5	0/1			
207	0/4				
214	0/1	0/2			
221				0/1	
226			0/1	0/3	
231			0/1	0/2	
233				0/1	
237			0/2	1/2	Cull Mammary tumour/ adenomas aged 233 days
241			1/1	0/2	Cull adenomas, aged 213 days
260			1/1	0/2	Cull adenomas aged 232 days
258			1/5	1/2	Tg+ Cull adenomas aged 226 days Tg- Cull adenomas aged 212 days
265				0/1	
269			0/1	0/1	
286			1 /2	0/1	Cull adenomas aged 281 days
335			1/1		Cull adenomas aged 331 days

Analysis of the data in table 13 suggests that the presence of the *APC* transgene does not result in the elimination or reduction in the frequency of the *Apc* $^{+/Min}$ phenotype being detected. A total of 33 *Apc*$^{+/Min}$ mice on a 87.5% C57BL/6 background were analysed using the Chi squared test to determine whether there was a significant difference between the number of mice developing adenomas with the prescence or abscence of the *APC* transgene. Five out of 15 transgene positive mice were culled due to adenomas whilst 2/18 transgene negative mice were culled. No significant difference was found (χ^2=2.418; p=0.12).

To determine whether the transgene could rescue the embryonic lethal phenotype associated with the homozygous Min genotype, transgene positive mice who were heterozygous for the Min allele were interbred. From such matings, one would expect 75% of offspring to be transgene positive; of which 25% would be homozygous for the *APC* wildtype allele, 50% heterozygous for the Min allele and 25% homozygous for the Min allele.

Forty eight mice were generated, one may expect to see 36 transgene positive mice of which 9 would be homozygous for the *Apc* wildtype allele; 18 heterozygous for the Min allele and 9 homozygous for the Min allele. In this study 38 mice were transgene positive, 26 were heterozygotes (*Apc*$^{+/Min}$) whilst 12 were homozygous for the *Apc* wildtype allele (*Apc*$^{+/+}$). No offspring were generated that were homozygous for the Min allele. Chi square test was used to determine whether there was a significant difference between the observed frequency of Tg+ Min homozygotes (0/48) versus the expected number of Tg+ Min homozygotes (9/48), results were statistically significant with χ^2=11.08; p<0.001. Results are summarised in table 14. It appears that the *APC* transgene is unable to rescue the embryonic lethal phenotype associated with the *Apc* $^{Min/Min}$ genotype.

Table 14: Analysis of Tg+; *Apc*$^{+/Min}$ crosses following the generation of 48 progeny

Genotype	Expected number (assuming no embryonic lethality)	Actual number
Tg+	36	38
Tg+; *Apc* $^{+/+}$	9	12
Tg+; *Apc* $^{+/Min}$	18	26
Tg+; *Apc* $^{Min/Min}$	9	0
Tg-	12	10
Tg-; *Apc* $^{+/+}$	3	5
Tg-; *Apc* $^{+/Min}$	6	5
Tg-; *Apc* $^{Min/Min}$	3	0

5.6.10 Analysis of embryos to determine whether the *APC* transgene results in partial rescue of the embryonic lethal phenotype

To determine whether the transgene could result in partial rescue, ie if transgene positive *Apc* $^{Min/Min}$ mice were viable for a longer period during development than that seen in *Apc* $^{Min/Min}$ (see section 5.1.1), matings were set up between Tg+; *Apc*$^{+/Min}$ mice. The date of coital activity was marked by the presence of a copulatory plug. At 8 dpc female mice were sacrificed by cervical dislocation and the uterus removed and placed into a petri dish containing PBS. Embryos were isolated from the uterus, numbered and placed in 70% ethanol. Using a dissection microscope, the yolk sac from each embryo was removed and placed into a sterile eppendorf. The placenta was discarded. All embryos within the litter were analysed by eye to determine whether any development abnormalities could be detected (carried out by Dr Alan Clarke). Following visual examination, the embryos were placed into 4% formalin prior to paraffin embedding.

DNA was extracted from the yolk sacs and subjected to PCR to determine the transgene status of the embryo. Methods have been described previously in sections 5.6.7.4 and 5.6.3.2.

5.6.10.1 Results

Twenty-one embryos were isolated from 4 females. Despite attempting to harvest embryos at 8 dpc, in some cases the embryos from litters were a couple of days older or younger. Two embryos were noted to exhibit exencephaly, the age of these embryos were day 9 and day 10. Eighteen embryos were transgene positive, of which 6 were *Apc* wildtype and 12 were *Apc* heterozygotes. No embryos homozygous for the Min mutation were detected. There was a statistically significant difference between the observed (0/21) and expected (4/21) frequency of *Apc* Min homozygotes, (χ^2= 4.941; $p<0.05$). This result indicates that the *APC* transgene is not partially rescuing the Min homozygous phenotype. The genotypes of all embryos and the expected number of embryos for each

genotype are shown in table 15. Both mice exhibiting exencephaly had the genotype Tg+; *Apc*$^{+/Min}$

Table 15: Genotypes of 21 embryos generated through mating Tg+; *Apc* $^{+/Min}$ mice.

Genotype	Expected number (assumming no embryonic lethality)	Actual number
Tg+	16	18
Tg+; *Apc* $^{+/+}$	4	6
Tg+; *Apc* $^{+/Min}$	8	12
Tg+; *Apc* $^{Min/Min}$	4	0
Tg-	5	3
Tg-; *Apc* $^{+/+}$	1	1
Tg-; *Apc* $^{+/Min}$	3	2
Tg-; *Apc* $^{Min/Min}$	1	0

5.7 Discussion

The *APC* transgene was successfully constructed and contained the *APC* cDNA sequence situated upstream of the SV40 poly A sequence. The expression of the transgene was driven by the promoter sequence of the murine *Pgk* gene; a gene that is ubiquitously expressed (McBurney *et al.*, 1991; 1994; Sutherland *et al.*, 1995). The *APC* cDNA, poly A and promoter sequences were flanked by loxP sites which lie in the same relative orientation in order to excise the intervening sequences following exposure to Cre recombinase (see section 5.1.2.1). Prior to injecting the transgene into oocytes for the production of transgenic mice, experiments were carried out to determine whether the DNA sequences, which lie between the two-loxP sites, could be excised in the presence of Cre-recombinase. This was carried out by transforming the transgene into a strain of E.Coli, which constitutively expresses Cre recombinase. Analysis of clones, which resulted from this transformation, showed that the DNA sequence between the loxP sequences could be efficiently excised from the transgene. No excision was seen following transformation of the incomplete transgene, which lacked one loxP site (see section 5.5.1).

The transgene was injected into the pronuclei of 110 oocytes; this resulted in the production of 13 pups. The number of pups born was a lot lower than had been expected, one would expect approximately one third of injected oocytes to result in live pups, therefore the expected number would be 39 (Dr A Mackenzie, personal communication). The low number observed may be an indication that a threshold exists where overexpression of *APC* results in embryonic lethality (discussed further below).

Of the 13 pups, six were lost due to water run outs within the cages, this occurred before the pups could be screened for the presence of the transgene. The remaining 7 pups were screened and two were shown to be transgene positive. One of the transgene positive mice was very weak and subsequently died aged 4 weeks. This mouse may have been a runt. Alternatively death may have been a result of the transgene integrating and disrupting a critical gene, or due to the level of *APC* expression.

There is evidence to suggest that overexpression of *APC* results in a phenotype in developing mouse embryos. Chimeric transgenic mice have previously been generated which express *APC* cDNA in intestinal and stomach epithelial cell lineages from embryonic day 15 (Wong *et al.*, 1996). Mice were viable and there was no significant difference in body weights between *APC* chimeras and normal mice. Analysis of the intestine epithelium revealed disordered cell migration. This phenotype was shown not to be caused by disruption in the levels of apoptosis or cellular proliferation, but was a result of disruption of cell adhesion, which resulted in the production of a non-adhesive migratory phenotype. It is still unclear how the APC protein regulates cell adhesion and cell migration.

Cell adhesion plays a critical role in embryonic and organ development and maintenance of tissue integrity (Takeichi 1995; Gumbiner 1996; Vleminckx 1999). The current understanding of Apc's role in cell adhesion is discussed in section 2.7. Briefly, APC can bind to free β-catenin, β-catenin that is not already in a stable complex with cadherin at the cell membrane and form a complex with Axin. This complex is phosphorylated by GSK-3β, and β-catenin is targeted for degradation. One may propose that increased levels of APC may result in a decrease in cell adhesion, resulting from increased β-catenin degradation and hence reduced β-catenin/cadherin complexes. However, in the study by Wong and co-workers overexpression of APC in chimeras did not result in a decrease in β-catenin levels.

In vitro studies have shown that the APC protein associates with microtubules and is localised in punctate clusters near the ends of microtubules that protrude into actively migrating membrane structures, this suggests that APC has a role in microtubule dependent cell migration (see section 2.3.4). The importance of APC in cell migration has recently been elucidated by the analysis of the Min model, mice who harbour a mutation at codon 850 (Mahmoud *et al.*, 1997). When compared to wildtype mice on the same genetic background, analysis of histologically normal intestinal epithelium exhibited elevated β-catenin expression which was associated with decreased proliferation, decreased apoptosis and a decreased rate of enterocyte crypt-villus migration (Mahmoud

et al., 1997).

It is likely that the level of *APC* expression will influence whether transgenic mice are viable. As the amount of *APC* expressed within each injected oocyte will vary, due to the transgene copy number, and to some extent, on the position of integration, then some mice may be viable and some may not. The level of expression may be such that a live pup is born but cannot survive.

To determine the transgene copy number in the surviving mouse, Southern blot protocol was used. However, despite numerous attempts, the sensitivity of the procedure could not be increased sufficiently to detect less than 100 copies of the transgene per single copy of murine genome. The amount of genomic DNA analysed was increased from 10μg up to 60 μg, this a relatively large amount of DNA to be used for this protocol as it has been reported that a single-copy of a gene can be detected in 10μg of human DNA (Old and Primrose 1989). It is possible that during the extraction, the DNA was excessively sheared or may have been degraded so that large molecular weight fragments were not present. The *APC* transgene is a relatively large fragment of 13.2 kb. If this were the case then, following incubation of the DNA sample with the restriction endonuclease and subsequent electrophoresis through the agarose gel, the majority of the DNA would appear as low molecular weight DNA. This was not the case as an intense streak of DNA was seen along side the full length of the molecular weight markers. This result also indicated that the DNA was being successfully transferred onto the Nylon membrane. Future experiments need to be carried out to increase the sensitivity of this protocol, one may use a probe targeted at a different region of the transgene, or the hybridisation conditions and stringency washes could be varied.

The failure to detect the *APC* transgene by Southern blot protocol also meant that I could not determine whether progeny generated during the breeding program were homozygous or heterozygous for the transgene. If the Southern blot protocol had been successful, mice heterozygous for the transgene could have been detected by a two-fold increase in the intensity of a 13.2 kb band detected by autoradiograph when compared to the heterozygous mouse DNA.

The transgene positive founder mouse was enrolled in a breeding program to establish transgene positive *Apc* Min homozygotes (Tg+; $Apc^{Min/Min}$) on a pure bred C57BL/6 background. Therefore, mating was carried out with the Min mouse ($Apc^{+/Min}$). During the program thirty-two Tg+; $Apc^{+/+}$ mice were generated. Three of these mice were culled to determine whether RNA was being transcribed from the transgene. RNA was extracted from various tissues (tissues are listed in figure 20) and subjected to reverse transcription PCR. In all tissues tested (see figure 20), transgene specific RNA was detected, all negative controls were negative. One would expect to see expression within all tissues as the promoter sequence, derived from the *Pgk* gene, is ubiquitously expressed (McBurney *et al.*, 1991; 1994; Sutherland *et al.*, 1995). This experiment was not designed to determine the level of expression within each tissue, but this could be included in later experiments.

To determine the presence of transgene specific protein, immunohistochemistry may be carried out using an antibody that recognises the human APC protein but not murine Apc protein. One such antibody is a commercially available (Oncogene research product, MA), in future studies this antibody could be optimised and used for the detection of the protein transcribed from the *APC* transgene.

The three mice, which were culled to detect RNA production and the remaining 25 mice which overexpress *APC*, whose ages range from 178 to 305 days, appear phenotypically normal. As a previous study has shown that forced expression of *APC* within the intestine results in disordered migration of cells up the colonic crypt (Wong *et al.*, 1996), one may expect a similar phenotype in mice from this study. Initial analysis of tissue sections taken from three mice with this genotype (Tg+; $Apc^{+/+}$), showed no evidence of disordered migration within the colon. No histological changes were noted in any of the other tissues analysed (see figure 20 for tissues analysed). The transgene chimeric mice generated by Wong and co-workers, expressed *APC* cDNA under the control of a fatty acid binding protein gene promoter (*Fabpl*). The promoter has been extensively characterised in transgenic and chimeric transgenic animals and is shown to direct gene expression along the entire length of the crypt-villus (Kim *et al.*, 1993; Hermiston and Gordon 1995;

Hermiston *et al.*, 1996). Expression is initiated in proliferating crypt epithelial cells and is sustained as all lineages complete their migration-associated differentiation (Trahair *et al.*, 1989). Conversely, the Pgk promoter sequence is known to drive a high level of expression in proliferating cells and a lower level in more differentiated cells, therefore the level of expression is likely to decrease from crypt to villus. The difference in promoter activity between these two models may result in phenotypic differences. To conclude whether a phenotype is associated with expression of the *APC* transgene in this study more mice need to be analysed. Furthermore, as these mice were heterozygous for the transgene, homozygous mice need to be examined to determine whether in cases where twice as much APC protein is expressed (transgene specific) a phenotype is exhibited.

Mice on a pure bred C57BL/6 background, which are heterozygous for the Min allele, develop multiple colonic adenomas (average number 50) at approximately 5 months (155 days) of age (see section 5.1.1). Symptoms of multiple adenomas, include anaemia (as determined by white feet), blood in the faeces and ultimately blockage of the colon. One aim of this study was to determine whether the presence of the *APC* transgene could reduce the number of adenomas and hence delay the onset or even completely eliminate this phenotype. As the age at which adenomas develop and the number of adenomas present per mouse has been shown to be dependent on the genetic background (see section 5.1.1) the data collected must be analysed according to the genetic background of the progeny, in this case either being 75% or 87.5% C57BL/6.

To date, no mice on a 75% C57BL/6 background were culled due to illness. The age of these mice ranged from 177-214 days and consisted of 23 Tg+; $Apc^{+/Min}$ mice and 3 Tg-; $Apc^{+/Min}$ mice. The low number of Tg negative mice prevents me from commenting on whether, in this group of animals, the transgene is having any effect.

Thirty-three mice on an 87.5% C57BL/6 background were analysed (see table 13), this group consisted of 15 mice with the genotype Tg+ $Apc^{+/Min}$ and 18 with the genotype Tg-; $Apc^{+/Min}$. To date, 7 mice have been culled due to illness, all exhibited symptoms of colonic adenomas and adenomas were detected in the small intestine following dissection.

The number of adenomas per mouse has not yet been determined. Statistical analysis has shown that there is no significant difference between the number of Tg+; $Apc^{+/Min}$ mice (5/15) and Tg-;$Apc^{+/Min}$ mice (2/18) developing ademomas (χ^2=2.418; p=0.12) These data therefore suggest that the transgene is not able to rescue the $Apc^{+/Min}$ phenotype. The average age of death of transgene negative mice was 222 days, whilst for transgene positive animals it was 257 days. As expected, the age of death is greater than that seen on a pure C57BL/6 background.

Elimination or reduction in the severity of the phenotype may only occur in transgene homozygotes, in this study progeny were derived either from interbreeding $Tg^{+/-}$; $Apc^{+/Min}$ genotypes or $Tg^{+/-}$; $Apc^{+/Min}$ with Tg-/-; $Apc^{+/Min}$ mice. Therefore, the number of progeny which would be homozygous for the transgene would be less than 25% (estimated at <8). During the breeding program, 48 transgene positive mice were generated from interbreeding $Tg^{+/-}$; $Apc^{+/Min}$ mice. The relative percentage of each possible genotype from these offspring, in the absence of embryonic lethality, would be expect to be 25% $Apc^{+/+}$; 50% $Apc^{+/Min}$ and 25% $Apc^{Min/Min}$. Twenty-five percent (12/48) were homozygous for the wildtype allele ($Apc^{+/+}$), 54% (26/48) were heterozygotes ($Apc^{+/Min}$). No offspring were generated that were homozygous for the Min allele. One may expect to see approximately 9 progeny with this genotype generated from such matings. Using a chi square test the difference between the expected and observed frequency is statistically significant (χ^2=11.08; p<0.001). Therefore, it appears that the *APC* transgene is unable to rescue the embryonic lethal phenotype associated with the $Apc^{Min/Min}$ genotype. However, it may be necessary for the offspring to be homozygous for the transgene. One would expect 25% of progeny from these matings to have the genotype $Tg^{+/+}$, of which, 25% would be homozygous for the Min allele. This equates to 12 $Tg^{+/+}$ mice, 3 of which would be predicted to be homozygous for the Min allele. In conclusion, more mice need to be generated to say with confidence that homozygosity for the transgene does not rescue the embryonic lethal Min phenotype.

In previous studies embryos homozygous for the *Apc* Min allele show development abnormalities at 6.5 days post coitum (dpc), embryos fail to develop the primitive

ectoderm and die before day 8 gestation (see section 5.1.1). By mid gestation (approximately 10 dpc) homozygotes consist of a mass of trophoblast giant cells with an additional cluster of much smaller embryonic cells. To determine whether the transgene could result in a partial rescue of the embryonic phenotype, i.e. whether expression could increase embryonic viability, Tg+ *Apc*$^{+/Min}$ crosses were set up and the embryos harvested at 8 dpc. Twenty-one embryos were isolated from four females. Despite attempting to harvest embryos at 8 dpc, in some cases the embryos from litters were a couple of days older or younger. Eighteen out of the 21 embryos were transgene positive, of which 12 were *Apc* heterozygotes and six were *Apc* wildtype. No embryos homozygous for the Min mutation were detected. One may predict from such mating, 5 embryos would be homozygous for the Min allele. Using statistics to analyse the observed number of mice (0/21) compared to the expected number mice (4/21) with this genotype, this result is statistically significant (χ^2=4.941; $p<0.05$). These data suggest that the APC transgene is not able to partially rescue the Min embryonic lethal phenotype.

Two embryos were shown to exhibit exencephaly, both of which had the genotype Tg+; *Apc*$^{+/Min}$. The age of these embryos was day 9 and day 10. Exencephaly has been noted to occur in a low percentage of embryos generated from *Apc* $^{+/Min}$ crosses (Dr J. Armstrong personal communication). It is therefore unclear whether this phenotype is a result of the presence of the transgene. To determine whether this is the case, more embryos need to be analysed and the relative frequency of exencephaly between transgene positive heterozygotes and transgene negative heterozygotes determined.

The breeding program will be continued to increase the numbers of progeny which overexpress *APC* and the number of transgene positive and negative *Apc* heterozygotes. Ideally, all mice will be on the pure bred C57BL/6 background so that comparisons between the phenotypes exhibited by different genotypes can be made without the influence of modifier loci (see section 5.1.1). More embryos will be harvested to determine whether the transgene can result in partial rescue of the *Apc* $^{Min/Min}$ embryonic lethal phenotype and to conclude whether exencephaly is associated with the presence of the *APC* transgene.

Once a large colony of mice has been established and the mice fully characterised, experiments may be carried out to excise the transgene. Ideally, I hoped to see embryonic rescue or partial rescue of the homozygous Min phenotype, and subsequently the *APC* transgene could be excised in a temporal or tissue specific manner, and the role of *APC* in development or the adult tissue further elucidated. The *Cre* gene could be delivered using the replication deficient recombinant adenovirus (described in section 5.1.2.2b and 5.6.8), resulting in global or tissue specific *APC* excision, or following breeding with *Cre* transgenic mice.

Within this department several transgenic mice which expresses *Cre* are available or are currently being developed. These include mice who express Cre specifically in lung tissue (under the control of a lung specific promoter), mammary tissue (under the control of the β-lactoglobulin promoter, Selbert *et al.*, 1997) and transgenic mice which express *Cre* specifically in the intestine and whose expression can be controlled in a temporal manner. In the latter case, expression is driven by the *Cdx-1* promoter sequence (Meyer and Gruss, 1993). To confer temporal control of Cre expression, Cre is being expressed as a fusion protein with a mutant oestrogen receptor ligand-binding domain (see section 5.1.2.2a). The mutated oestrogen receptor ligand-binding domain is insensitive to the endogenous hormone β-oestradiol but still responsive to the synthetic oestrogen antagonist 4-OH-tamoxifen (Schwenk *et al.*, 1998).

Therefore, this *APC* transgenic mouse model has a great deal of potential. However, to date it has not been established whether a protein product is being produced from the transgene. Immunohistochemical studies are of high priority. It is possible that during *in vitro* manipulation of the *APC* cDNA, a mutation within the cDNA may have occurred. If it was such that this resulted in the production of a premature translation stop codon, then a truncated product may be produced. The human specific APC antibody which would be used (APC Ab-4, Oncogene research products) is raised to the carboxyl terminus of the protein and hence protein would not be detected. This antibody is not suitable for western blotting techniques so a truncated product could not be detected by this means. If no protein were detected following immunohistochemical analysis then the

next step would be to sequence the transgene. This is a very large sequence and would be expensive and time consuming.

Another potential problem associated with this model is that human *APC* cDNA was used. Despite high homology between species (see section 2.5), this may result in a phenotype. Secondly, numerous transcripts have been shown to be produced from the *APC* gene as a result of alternative splicing (reviewed in section 2.2). Since the isolation of the *APC* cDNA sequence, which was used to generate this transgenic model (Groden *et al.*, 1991), four exons have been identified which are situated upstream of the original exon 1. These exons are not present within the cDNA sequence. Splice variants generated from these additional exons have been shown to be present at elevated levels in postmitotic and terminally differentiated tissues (Santoro and Groden 1997) suggesting that different APC isoforms may be involved in the physiology of cell cycle cessation. In this model, no splice variants are produced and all *APC* exons (1-15) are transcribed and translated, therefore,if different isoforms of APC specify various cellular functions, these specific functions may not be carried out in this model.

Since the initiation of this project, Shibata and co-workers have published details of a floxed *Apc* transgenic mouse. Exon 14 of the *Apc* gene has been floxed and mice generated which are homozygous for this feature (Shibata *et al.*, 1997). Delivery of Cre recombinase to these mice resulted in deletion of the loxP flanked exon 14 and also the introduction of a frame-shift mutation at codon 580. Protein produced from this gene would contain the homodimerisation domain and part of the armadillo repeats (see section 2.3), all downstream domains would be lost. Adenovirus encoding the Cre recombinase gene was delivered via enema to the intestine and resulted in the deletion of *Apc* exon 14 within the large intestine. Unfortunately, this delivery method was inefficient and only a small area of the intestine underwent Cre mediated excision. Within this area, the mice developed adenomas within approximately 4 weeks. This model will allow extensive analysis of the function of APC in both embryogenesis and the adult tissue.

Chapter 6: Summary

APC is well established as a tumour suppressor gene and inactivation of the gene is associated with numerous types of cancer. Germline mutations within the *APC* gene predispose individuals to FAP, a disease whereby multiple adenomas develop within the colon predisposing the individual to colorectal cancer and in some instances extracolonic manifestations (Bodmer *et al*, 1987; Leppert *et al.*, 1987; Dunlop *et al.*, 1990). Analysis of subsequent tumours reveals mutation or loss of the remaining *APC* allele (Nishisho *et al.*, 1991; Groden *et al.*, 1991). Biallelic inactivation of the *APC* gene has also been reported in sporadic colorectal, gastric and hepatic cancers (Nishisho *et al.*, 1991; Groden *et al.*, 1991;Tamura *et al.*, 1994; Oda *et al.*, 1996; Imai *et al.*, 1997). Mutations within the *APC* gene have also been detected in pancreatic, oesophageal and prostate cancers (Yashima *et al.*, 1994; Horri *et al.*, 1992a; Powell *et al.*, 1994; Watanabe *et al.*, 1996; Gonzalez *et al.*, 1997). It is currently unclear the precise role APC plays in carcinogenesis but the APC protein appears to have multiple functions within the cell, including the regulation of cell proliferation and differentiation, cell adhesion and signalling, and cell movement.

The human *APC* gene is located at chromosome 5q21 (Kinzler *et al.*, 1991a,b; Joslyn *et al.*, 1991). Several groups have reported genetic loss at this region within NSCLC (Ashton-Rickardt *et al.*, 1991; D'Amico *et al.*, 1992; Tsuchiya *et al.*, 1992; Horri *et al.*, 1992a; Hosoe *et al.*, 1994; Wieland and Bohm 1994, Fong *et al.*, 1995). The frequency of loss varies between studies from 20-71%. The varying frequency of LOH may be a reflection of different histological types within cohorts as a number of studies have revealed that ADC and SCC exhibit different allelotypes. Furthermore varying frequencies of LOH may reflect tumour stage.

Through the analysis of polymorphic sites within the *APC* gene and a second gene lying approximately within 500 kb downstream, the *MCC* (mutated in colorectal cancer) gene, I have determined the frequency of loss of heterozygosity (LOH) in a cohort of consecutive pulmonary ADC and SCC from Edinburgh Royal Infirmary. Loss of

heterozygosity was shown to occur in 46% (11/24) of SCC and 36% (16/44) of ADC. There was no significant difference between frequency of loss in these two histological subtypes. Analysis of tumours according to tumour stage, revealed that LOH was not associated with increasing tumour stage in SCC or ADC. Therefore, this result indicated that LOH at 5q21 is not associated with tumour progression in ADC or SCC.

Subclassification of ADC into the site of origin; bronchial or parenchymal, revealed no significant difference in the frequency of LOH between these two subgroups or between the different tumour stages within each subgroup.

There are currently no studies that identify the point at which LOH at 5q21 occurs in NSCLC. As several stages of premalignant lesions have been identified in the development of SCC, analysis would be relatively easy; however, such a study would not be so easy to conduct for ADC as premalignant lesions have been less well characterised. There is now extensive evidence to suggest that alveolar atypical hyperplastic lesions are precursor lesions to parenchymal ADC. These lesions may be studied to determine the frequency of LOH at 5q21 and the frequency compared to that seen in parenchymal ADC. No premalignant lesions for bronchial ADC have been identified.

In the study described above, polymorphic sites within two genes, *APC* and *MCC* were analysed. In all cases that were informative for both genes; loss or retention was always seen in both genes, i.e. there were no cases showing discordance between the two genes, indicating that this was a minimum area of loss. In previous studies of NSCLC, a common region of chromosomal deletion at 5q21 has been identified. This region spans 3-5 megabases and harbours numerous genes including the tumour suppressor gene *APC* and the *MCC* gene. To date, of the genes identified and characterised within this region, the most likely candidate gene to be involved in NSCLC tumourigenesis is *APC*.

According to Knudson's hypothesis, tumour suppressor genes act recessively at the cellular level so that both copies of the gene must be inactivated for growth suppressive functions to be eliminated (Knudson 1978). Therefore, demonstrating the loss of one *APC* allele in a proportion of tumours would not indicate loss of tumour suppression. The final aim of this initial investigation was to determine whether biallelic inactivation of the

APC gene is a characteristic of pulmonary ADC and SCC.

Extensive studies of several types of cancer have demonstrated that the majority of mutations within the *APC* gene occur between codon 1286-1513, named the mutation cluster region, and that these mutations nearly always result in the production of a truncated protein product (See Laurent-Puig *et al.*, 1998).

In cases of NSCLC that exhibited LOH at 5q21, the mutation cluster region was examined for mutations. This was carried out using single strand conformational polymorphism analysis. No mutations were detected within the MCR of 30 tumours analysed. It is possible that mutations exist outside this region. As the majority of mutations within this gene have been shown to result in the truncation of the protein product, immunohistochemistry was carried out using an antibody raised to the carboxyl terminus of the APC protein. In all cases analysed (N=30), full length APC was shown to be present and the cytoplasmic and nuclear staining seen in normal tissue was also detected in the tumours. Therefore, using SSCP and immunohistochemistry there was no evidence of biallelic inactivation of the *APC* gene. One could state that mutations within the mutation cluster region are extremely rare in pulmonary ADC and SCC. Furthermore, if mutations occur in the *APC* gene they rarely result in the production of truncated APC protein. The detection of full length APC in these cases also suggests that epigenetic factors do not lead to loss of *APC* expression in these tumours. If the *APC* gene did play a role in lung carcinogenesis, one would expect individuals with germline mutations in the *APC* gene to have an increased risk of lung cancer and this is not the case.

The proto-oncogene *K-RAS* encodes a protein that functions as a guanosine diphosphate /guanosine triphosphate (GDP/GTP) regulated switch, transducing extracellular stimuli to cytoplasmic signal transduction cascades. The *K-RAS* gene can acquire transforming potential following point mutations that lead to amino acid changes at codons 12, 13 or 61 or through mutations at or near the GTP-binding domain, both of which result in enhanced ability to retain GTP leading to continuous signalling (Fayed and O'Brien, 1995).

The *K-RAS* gene is mutated in 25-50% of ADC and the majority of mutations occur at codon 12 (>90%) (Rodenhuis *et al.*, 1988; Slebos *et al.*, 1989; Reynolds *et al.*, 1991, Mitsudomi *et al.*, 1991; Slebos *et al.*, 1991; Mills *et al.*, 1995; Keohauana *et al.*, 1996; Graziono *et al.*, 1999; Gealy *et al.*, 1999). Mutations at codons 13 and 61 are relatively rare. Mutations at codon 12 are generally guanine to thymine (G-T) transversions (approximately 80%), occurring at position 1 in 60% of tumours and position 2 in 15% of tumours. G-A transitions also occur, these are usually at position 2 (20% of tumours). Guanine to thymine transversions are typical of the production of DNA adducts at guanine residues and this mutation has been shown to be associated with the carcinogens derived from cigarette smoking.

Previous investigations have reported that *K-RAS* mutations are not associated with tumour progression, and that mutation within this gene may be an early event in the development of pulmonary ADC, with mutation occurring prior to clonal expansion (Li *et al.*,1994a). Due to the absence of recognised precursor lesions, it has been difficult to further define the timing of *K-RAS* mutations in the development of pulmonary ADC.

In this study, I determined that mutations at position 1 or 2 of codon 12 of the *K-RAS* gene occurred in 25% (16/65) of ADC in a cohort obtained from Edinburgh Royal infirmary. The majority of mutations were G-T transversions at position 1 (56%; 9/16) whilst other mutations occured at a lower frequency; G-T transversions at position 2 of codon 12 occured in 38% (6/16) of cases and G-A transition at position 2 occured in 6% (1/16) of cases. The presence of *K-RAS* mutation did not increase significantly with tumour stage. These data, derived from the analysis of this cohort are similar to results obtained with previously published studies.

Subclassification of ADC into bronchial and parenchymal origin showed that significantly more mutations were associated with ADC of parenchymal origin than of bronchial origin. This is the first reported genetic difference between these two histological subtypes and suggests that different genetic pathways are involved in the development of these tumours. It is interesting to note that like bronchial ADC, SCC and SCLC, that also arise in the central portion of the lung and are associated with smoking, are not frequently

characterised as having *K-RAS* mutations.

Analysis of 32 areas of alveolar atypical hyperplasia (AAH), possible precursor lesions for parenchymal ADC, identified *K-RAS* mutations. Mutations were identified in sixteen patients, 10 of which harboured an ADC, revealed the presence of a codon 12 mutation in 5 lesions (16%). Two positive lesions were identified in case 1, which contained 5 lesions plus an ADC. No mutations were detected in the ADC and different mutations were detected within the two AAH lesions (G-A transition pos.1 and a G-T transversion at position 1 and 2). In case 7, one out of 6 AAH lesions harboured a *K-RAS* codon 12 mutation, no mutation was detected in the corresponding ADC. In cases 13 and 15, which consisted of only one lesion and no ADC, mutations were also detected.

The mutation spectrum seen within AAH lesions consisted of 7 cases of G-T transversions and only one lesion with G-A transitions. This mutation spectrum is consistent with mutations detected in parenchymal ADC further strengthening the proposal that AAH lesions are precursors of parenchymal ADC. These data also suggest that assuming AAH lesions are true preneoplastic lesions, mutations within *K-RAS* gene are a very early event in the development of parenchymal ADC. Since the initiation of this work, two studies have reported the presence of *K-RAS* mutations within areas of AAH.

From the different genetic backgrounds of AAH lesions and corresponding ADC within patients analysed in this study, and the high percentage of G-T transversions, one may suggest that AAH lesions are independent entities, independent of ADC and of other lesions. These lesions may be a result of a phenomenon known as field carcinogenesis (Slaughter *et al.*, 1953); in this case the carcinogen is likely to be cigarette smoke. Alternatively, multiple AAH lesions may be a result of spread of progeny from a single progenitor cell and mutation of the *K-RAS* gene is an additional genetic event, this phenomenon has been referred to as subclonal drift (Nowell, 1976).

To further characterise the function of APC, an *APC* transgene was constructed. This transgene contains the human *APC* cDNA sequence whose expression is driven by the promoter sequence of the ubiquitously expressed murine phosphoglycerate kinase (*Pgk)* gene. These sequences are flanked by 2 loxP sites that lie in the same orientation thereby

enabling excision of the flanked sequence following exposure to Cre recombinase. Excision was demonstrated in a prokaryotic system.

Transgene positive mice have been generated following microinjection of the transgene into the pronuclei of oocytes. The mice generated contained the *APC* transgene on a murine wildtype *Apc* background (Tg+; *Apc* $^{+/+}$). Excision of the DNA sequence flanked by the loxP sites has been demonstrated in an eukaryotic system, using embryonic fibroblasts derived from transgene positive embryos.

Analysis of mice with the genotype Tg+; *Apc*$^{+/+}$ has shown that RNA is transcribed from the transgene. To date no phenotype has been noted in the mice that represent overexpression of *APC*.

Transgene positive mice have been bred with Min mice (*Apc* $^{+/Min}$); Min mice harbour a germline mutation that results in a truncated Apc protein at codon 850. Mice heterozygous for the truncated (*Min*) allele exhibit an FAP phenotype and develop numerous colonic adenomas. Homozygosity for the *Min* allele results in embryonic lethality. The aim of this part of the thesis was to determine whether the *APC* transgene could result in reduction in severity or elimination of the FAP phenotype exhibited by *Apc* $^{+/Min}$ mice. As the severity of the phenotype is modified by modifier loci, comparison of transgene positive and transgene negative *Apc* heterozygotes has been carried out between mice on the same genetic background. Until the colony has been established on a pure bred background, the numbers available for analysis are low. However, data from mice on the 87.5% C57BL/6 background have revealed that the *APC* transgene can not eliminate the development of multiple colonic adenomas and that there is no statistically significant differene in the frequency of the *Apc*$^{+/Min}$ phenotype in transgene positive and transgene negative mice (χ^2=2.418; p=0.12). It also appears that the embryonic lethal phenotype can not be eliminated by *APC* transgene as there was a significant difference between the observed frequency of Tg+ Min homozygotes and the expected frequency. Expression of the transgene may result in partial rescue of embryonic lethality. To date a small number of embryos have been analysed and no *Apc* $^{Min/Min}$ homozygotes have been detected, the results generated indicated that there is statistically significant difference in

the observed number of min homozyotes and expected number of Min homozygotes, data which indicates that the *APC* transgene is not able to result in partial rescue of the Min homozygote lethal phenotype. Two embryos were identified that exhibited exencephaly and both embryos had the genotype Tg+; *Apc* $^{+/Min}$. More embryos must be studied to determine whether this phenotype is associated with *APC* transgene expression.

In these studies, the majority of mice and embryos generated are heterozygous for the *APC* transgene, further investigations need to be conducted to ascertain whether homozygosity for the transgene can firstly eliminate or reduce the *Apc* Min heterozygous phenotype. Secondly, if homozygosity for the transgene results in rescue or partial rescue of the *Apc* $^{Min/Min}$ embryonic lethal phenotype.

References

Abdel-Rahman SZ, El-Zein RA, Zwischenberg JB, Au WW. Association of the NAT1*10 genotype with increased chromosome aberrations and higher lung cancer risk in cigarette smokers. Mutat Res 1998, **398**:43-45

Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. beta-catenin is a target for the ubiquitin-proteasome pathway. EMBO J 1997, **16**:3797-3804

Abremski K, Hoess R, Sternberg N. Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. Cell 1983, **32**:1301-1311

Agah R, Frenkel PA, French BA, Michael LH, Overbeek PA, Schneider MD. Gene recombination in postmitotic cells. Targeted expression of Cre recombinase provokes cardiac-restricted, site-specific rearrangement in adult ventricular muscle in vivo. J Clin Invest 1997, **100**:169-179

Aghib DF, McCrea PD. The E-cadherin complex contains the src substrate p120. Exp Cell Res 1995, **218**:359-369

Aizawa H, Emori Y, Murofushi H, Kawasaki H, Sakai H, Suzuki K. Molecular cloning of a ubiquitously distributed microtubule-associated protein with Mr 190,000. Biol Chem 1990, **265**:13849-13855

Akasaka K, Tamada M, Wang F, Kariya K, Shima F, Kikuchi A, Yamamoto M, Shirouzu M, Yokoyama S, Kataoka T. Differential structural requirements for interaction of Ras protein with its distinct downstream effectors. J Biol Chem 1996, **271**:5353-5360

Alavanja MC, Brownson RC, Benichou J. Estimating the effect of dietary fat on the risk of lung cancer in nonsmoking women. Lung Cancer 1996, **14** Suppl 1:S63-74

Alderson MR, Rattan NS, Bidstrup L. Health of workmen in the chromate-producing industry in Britain. Br J Ind Med 1981, **38**:117-124

Alitalo K, Schwab M. Oncogene amplification in tumour cells. Adv Cancer Res 1986, **47**:235-281

Almog N, Rotter V. An insight into the life of p53: a protein coping with many functions! Review of the 9th p53 Workshop, Crete, May 9-13, 1998. Biochim Biophys Acta 1998, **1378**:R43-54

Alroy I, Yarden Y. The ErbB signaling network in embryogenesis and oncogenesis:signal diversification through combinatorial ligand-receptor interactions. FEBS Lett 1997, **410**:83-86

Amati B, Littlewood TD, Evan GI, Land H. The c-Myc protein induces cell cycle progression and apoptosis through dimerization with Max. EMBO J 1993, **12**:5083-5087

D' Amico D, Carbone DP, Johnson BE, Meltzer SJ, Minna JD. Polymorphic sites within the MCC and APC loci reveal very frequent loss of heterozygosity in human small cell lung cancer. Cancer 1992, **52**:1996-1999

Angres B, Barth A, Nelson WJ. Mechanism for transition from initial to stable cell-cell adhesion: kinetic analysis of E-cadherin-mediated adhesion using a quantitative adhesion assay. J Cell Biol 1996, **134**:549-557

Anton M, Graham FL. Site-specific recombination mediated by an adenovirus vector expressing the Cre recombinase protein: a molecular switch for control of gene expression. J Virol 1995, **69**:4600-4606

Apolinario RM, van der Valk P, deJong JS, Deville W, van Ark-Otte J, Dingemans AM, van Mourik JC, Postmus PE, Pinedo HM, Giaccone G. Prognostic value of the expression of p53, bcl-2, and bax oncoproteins, and neovascularization in patients with radically resected non-small-cell-lung cancer. J Clin Oncol 1997, **15**:2456-2466

Archer VE. Lung cancer risks of underground miners: cohort and case-control studies. Yale J Biol Med 1988, **61**:183-193

Armstrong RN. Structure, catalytic mechanism, and evolution of the glutathione transferases. Chem Res Toxicol 1997,**10**:2-18

Ashton-Rickardt PG, Wyllie AH, Bird BB, Dunlop MG, Steel CM, Morris RG, Piris J, Pomanowski P, Wood R, White R, Nakamura Y. MCC, a candidate familial polyposis gene in 5q21 shows frequent allele loss in colorectal and lung cancer. Oncogene 1991, **6**:1881-1886

Auer H, Warncke K, Nowak D, Koops F, Jaeger U, Kurzbauer R, Ruediger HW. Variations of p53 in cultured fibroblasts of patients with lung cancer who have a presumed genetic predisposition. Am J Clin Oncol 1999, **22**:278-282

Auerbach O, Gere JB, Pawlowski JM, Muehsam GE, Smolin HJ, Stout AP. Carcinoma in situ and early invasive carcinoma occuring in the tracheobronchial trees in cases of bronchial carcinoma. J. Thorac Surg 1957, **34**:298-309

Auerbach O, Stout AP, Hammond EC, Garfinkel L. Changes in bronchial epithelium in relation to cigarette smoking and in relation to lung cancer. N Engl J Med 1961, **265**: 253-267

Austin S, Ziese M, Sternberg N. A novel role for site-specific recombination in maintenance of bacterial replicons. Cell 1981, **25**:729-736

Bafico A, Gazit A, Pramila T, Finch PW, Yaniv A, Aaronson SA. Interaction of Frizzled Related Protein (FRP) with wnt Ligands and the Frizzled Receptor Suggests Alternative Mechanisms for FRP Inhibition of wnt Signaling. J Biol Chem 1999, **274**:16180-16187

Balsara BR, Sonoda G, du Manoir S, Siegfried JM, Gabrielson E, Testa JR. Comparative genomic hybridisation analysis detects frequent, often high-level, overrepresentation of DNA sequences at 3q, 5p, 7p, and 8q in human non-small cell lung carcinomas. Cancer Res 1997, **57**:2116-2120

Barbacid M. ras genes. Annu Rev Biochem 1987, **56**:779-827

Barlow C, Schroeder M, Lekstrom-Himes J, Kylefjord H, Deng CX, Wynshaw-Boris A, Spiegelman BM, Xanthopoulos KG. Targeted expression of Cre recombinase to adipose tissue of transgenic mice directs adipose-specific excision of loxP-flanked gene segments. Nucleic Acids Res 1997, **25**:2543-2545

Barsky SH, Huang SJ, Bhuta S. The extracellular matrix of pulmonary scar carcinomas is suggestive of desmoplastic origin. Am J Pathol 1986, **124**:412-419

Bartek J, Bartkova J, Lukas J. The retinoblastoma protein pathway and the restriction point. Curr Opin Cell Biol 1996, **8**:805-814

Behrens J, Mareel MM, Van Roy FM, Birchmeier W. Dissecting tumor cell invasion: epithelial cells acquire invasive properties after the loss of uvomorulin-mediated cell-cell adhesion. J Cell Biol 1989, **108**:2435-2447

Behrens J, von Kries JP, Kuhl M, Bruhn L, Wedlich D, Grosschedl R, Birchmeier W. Functional interaction of beta-catenin with the transcription factor LEF-1. Nature 1996, **382**:638-642

Behrens J, Jerchow BA, Wurtele M, Grimm J, Asbrand C, Wirtz R, Kuhl M, Wedlich D, Birchmeier W. Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. Science 1998, **280**:596-599

Beinhauer JD, Hagan IM, Hegemann JH, Fleig U. Mal3, the fission yeast homologue of the human APC-interacting protein EB-1 is required for microtubule integrity and the maintenance of cell form. J Cell Biol 1997, **139**:717-728

Belinsky SA, Swafford DS, Finch GL, Mitchell CE, Kelly G, Hahn FF, Anderson MW, Nikula KJ. Alterations in the K-ras and p53 genes in rat lung tumors. Environ Health Perspect 1997, **105** Suppl 4:901-906

Bennett WP, Colby TV, Travis WD, Borkowski A, Jones RT, Lane DP, Metcalf RA, Samet JM, Takeshima Y, Gu JR, *et al.* p53 protein accumulates frequently in early bronchial neoplasia. Cancer Res 1993, **53**:4817-4822

Berrueta L, Kraeft SK, Tirnauer JS, Schuyler SC, Chen LB, Hill DE, Pellman D, Bierer BE. The adenomatous polyposis coli-binding protein EB1 is associated with cytoplasmic and spindle microtubules. Proc Natl Acad Sci U S A 1998, **95**:10596-601

Bertrand P, Rouillard D, Boulet A, Levalois C, Soussi T, Lopez BS. Increase of spontaneous intrachromosomal homologous recombination in mammalian cells expressing a mutant p53 protein. Oncogene 1997, **14**:1117-1122

Betticher DC, Heighway J, Hasleton PS, Altermatt HJ, Ryder WD, Cerny T, Thatcher N. Prognostic significance of CCND1 (cyclin D1) overexpression in primary resected non-small-cell lung cancer. Br J Cancer 1996, **73**:294-300

Bhanot P, Brink M, Samos CH, Hsieh JC, Wang Y, Macke JP, Andrew D, Nathans J, Nusse R. A new member of the frizzled family from Drosophila functions as a Wingless receptor. Nature 1996, **382**:225-230

Bhat RV, Baraban JM, Johnson RC, Eipper BA, Mains RE. High levels of expression of the tumor suppressor gene APC during development of the rat central nervous system. J Neurosci 1994, **14**:3059-3071

Bhattacharya G, Boman BM. Phosphorylation of the adenomatous polyposis coli protein and its possible regulatory effects in cells. Biochem Biophys Res Commun 1995, **208**:103-110

Bienz M, Tremml G. Domain of Ultrabithorax expression in Drosophila visceral mesoderm from autoregulation and exclusion. Nature 1988, **333**:576-578

Bienz M. Homeotic genes and positional signalling in the Drosophila viscera. Trends Genet 1994, **10**:22-26

van Biesen T, Hawes BE, Luttrell DK, Krueger KM, Touhara K, Porfiri E, Sakaue M, Luttrell LM, Lefkowitz RJ. Receptor-tyrosine-kinase- and G beta gamma-mediated MAP kinase activation by a common signalling pathway. Nature 1995, **376**:781-784

Bilger A, Shoemaker AR, Gould KA, Dove WF. Manipulation of the mouse germline in the study of Min-induced neoplasia. Semin Cancer Biol 1996, 7:249-260

Birchmeier W, Behrens J. Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. Biochim Biophys Acta 1994, **1198**:11-26

Blackwell TK, Kretzner L, Blackwood EM, Eisenman RN, Weintraub H. Sequence-specific DNA binding by the c-Myc protein. Science 1990, **250**:1149-1151

Blackwood EM, Eisenman RN. MAX: A helix-loop-helix zipper protein that forms a sequence specific DNA binding complex with myc. Science 1991, **251**:1211-1217

Bodmer WF, Bailey CJ, Bodmer J, Bussey HJ, Ellis A, Gorman P, Lucibello FC, Murday VA, Rider SH, Scambler P, et al. Localization of the gene for familial adenomatous polyposis on chromosome 5. Nature 1987, **328**:614-616

Bolen JW, Thorning D. Histogenic classification of pulmonary carcinomas. Peripheral adenocarcinomas studied by light microscopy, histochemistry and electron microscopy. Pathol Annu 1982, **17**:17-100

Bollag G, McCormick F. Differential regulation of rasGAP and neurofibromatosis gene product activities. Nature 1991, **351**:576-579

Boriack-Sjodin PA, Margarit SM, Bar-Sagi D, Kuriyan J. The structural basis of the activation of Ras by Sos. Nature 1998, **394**:337-343

Bos JL. ras oncogenes in human cancer: a review. Cancer Res 1989, **49**:4682-4689

Bos JL. Ras-like GTPases. Biochim Biophys Acta 1997, **1333**:M19-31
Bouchardy C, Benhamou S, Dayer P. The effect of tobacco on lung cancer risk depends on CYP2D6 activity. Cancer Res 1996, **56**:251-253

Bouchardy C, Mitrunen K, Wikman H, Husgafvel-Pursiainen K, Dayer P, Benhamou S, Hirvonen A. N-acetyltransferase NAT1 and NAT2 genotypes and lung cancer risk. Parmacogenetics 1998, **8**:291-298

Bourne HR, Sanders DA, McCormick F. The GTPase superfamily: a conserved switch for diverse cell functions. Nature 1990, **348**:125-132

Bourne HR, Sanders DA, McCormick F. The GTPase superfamily: conserved structure and molecular mechanism. Nature 1991, **349**:117-127

Bourne HR. Colon cancer. Consider the coiled coil.... Nature 1991a, **351**:188-190

Bourne HR. Colon cancer. Suppression with a difference. Nature 1991b, **353**:696-697

Boynton RF, Blount PL, Yin J, Brown VL, Huang Y, Tong Y, McDaniel T, Newkirk C, Resau JH, Raskind WH, et al. Loss of heterozygosity involving the APC and MCC genetic loci occurs in the majority of human esophageal cancers. Proc Natl Acad Sci U S A 1992, **89**:3385-3388

Bradley RS, Brown AM. A soluble form of Wnt-1 protein with mitogenic activity on mammary epithelial cells. Mol Cell Biol 1995, **15**:4616-4622

Brannon M, Gomperts M, Sumoy L, Moon RT, Kimelman D. A beta-catenin/XTcf-3 complex binds to the siamois promoter to regulate dorsal axis specification in Xenopus. Genes Dev 1997, **11**:2359-2370

Brenman JE, Bredt DS. Synaptic signaling by nitric oxide. Curr Opin Neurobiol 1997, 7:374-378

Brinkmann V, Foroutan H, Sachs M, Weidner KM, Birchmeier W. Hepatocyte growth factor/scatter factor induces a variety of tissue-specific morphogenic programs in epithelial cells. J Cell Biol 1995, **131**(6 Pt 1):1573-1586

Brody SL, Crystal RG. Adenovirus-mediated in vivo gene transfer. Ann N Y Acad Sci 1994, **716**:90-101; discussion 101-3

Brown AL Prediction in the face of uncertainty: the United States versus Reserve Mining Company. Regul Toxicol Pharmacol 1986, **6**:75-79

Brunet A, Bonni A, Zigmond MJ, Lin MZ, Juo P, Hu LS, Anderson MJ, Arden KC, Blenis J, Greenberg ME. Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. Cell 1999, **96**:857-868

Bruning JC, Michael MD, Winnay JN, Hayashi T, Horsch D, Accili D, Goodyear LJ, Kahn CR A muscle-specific insulin receptor knockout exhibits features of the metabolic syndrome of NIDDM without altering glucose tolerance. Mol Cell 1998, **2**:559-569

Brunner E, Peter O, Schweizer L, Basler K. pangolin encodes a Lef-1 homologue that acts downstream of rmadillo to transduce the Wingless signal in Drosophila. Nature 1997, **385**:829-833

Bunyan DJ, Shea-Simonds J, Reck AC, Finnis D, Eccles DM. Genotype-phenotype correlations of new causative APC gene mutations in patients with familial adenomatous polyposis. J Med Genet 1995, **32**:728-731

Buckley MF, Sweeney KJ, Hamilton JA, Sini RL, Manning DL, Nicholson RI, deFazio A, Watts CK, Musgrove EA, Sutherland RL Expression and amplification of cyclin genes in human breast cancer. Oncogene 1993, **8**:2127-2133

Burchell B, Coughtrie MW. Genetic and environmental factors associated with variation of human xenobiotic glucuronidation and sulfation. Environ Health Perspect 1997,**105** Suppl 4:739-747

Burgering BM, Coffer PJ. Protein kinase B (c-Akt) in phosphatidylinositol-3-OH kinase signal transduction. Nature 1995, **376**:599-602

Butz S, Stappert J, Weissig H, Kemler R. Plakoglobin and beta-catenin: distinct but closely related. Science 1992, **257**:1142-1144

Cadigan KM, Nusse R. wingless signaling in the Drosophila eye and embryonic epidermis. Development 1996, **122**:2801-2812

Cagle PT, Fraire AE, Greenberg SD, Cox A, Brown RW. Potential utility of p53 immunopositivity in differentiation of adenocarcinomas from reactive epithelial atypias of the lung. Hum Pathol 1996, **27**:1198-1203

Cahill DP, Lengauer C, Yu J, Riggins GJ, Willson JK, Markowitz SD, Kinzler KW, Vogelstein B. Mutations of mitotic checkpoint genes in human cancers. Nature 1998, **392**:300-303

Campbell SL, Khosravi-Far R, Rossman KL, Clark GJ, Der CJ. Increasing complxity of Ras signaling. Oncogene 1998, **17**:1395-1413

Candelora EC, Stockwell HG, Armstrong AW, Pinkham PA. Dietary intake and risk of lung cancer in women who never smoked. Nutr Cancer 1992;**17**:263-270

Cardone MH, Roy N, Stennicke HR, Salvesen GS, Franke TF, Stanbridge E, Frisch S, Reed JC. Regulation of cell death protease caspase-9 by phosphorylation. Science 1998, **282**:1318-1321

Carey FA, Wallace WAH, Fergusson RJ, Kerr KM, Lamb D. Alveolar atypical hyperplasia in association with primary pulmonary adenocarcinoma: a clinicopathological study of 10 cases. Thorax 1992, **47**:1041-1043

Carlsson P, Waterman ML, Jones KA. The hLEF/TCF-1 alpha HMG protein contains a context-dependent transcriptional activation domain that induces the TCR alpha enhancer in T cells. Genes Dev 1993, 7:2418-2430

Carraway KL 3rd, Soltoff SP, Diamonti AJ, Cantley LC. Heregulin stimulates mitogenesis and phosphatidylinositol 3-kinase in mouse fibroblasts transfected with erbB2/neu and erbB3. J Biol Chem 1995, **270**:7111-7116

Caspari R, Olschwang S, Friedl W, Mandl M, Boisson C, Boker T, Augustin A, Kadmon M, Moslein G, Thomas G, et al. Familial adenomatous polyposis: desmoid tumours and lack of ophthalmic lesions (CHRPE) associated with APC mutations beyond codon 1444. Hum Mol Genet 1995, **4**:337-340

Cavallo RA, Cox RT, Moline MM, Roose J, Polevoy GA, Clevers H, Peifer M, Bejsovec A. Drosophila Tcf and Groucho interact to repress Wingless signalling activity. Nature 1998, **395**:604-608

Chan SD, Karpf DB, Fowlkes ME, Hooks M, Bradley MS, Vuong V, Bambino T, Liu MY, Arnaud CD, Strewler GJ, et al Two homologs of the Drosophila polarity gene frizzled (fz) are widely expressed in mammalian tissues. J Biol Chem 1992, **267**:25202-25207

Chen X, Bargonetti J, Prives C. p53, through p21 (WAF1/CIP1), induces cyclin D1 synthesis. Cancer Res 1995, **55**:4257-4263

Chen H, Paradies NE, Fedor-Chaiken M, Brackenbury R. E-cadherin mediates adhesion and suppresses cell motility via distinct mechanisms. J Cell Sci 1997, **110** :345-356

Chiba I, Takahashi T, Nau MM, D'Amico D, Curiel DT, Mitsudomi T, Buchhagen DL, Carbone D, Piantadosi S, Koga H *et al.* Mutations in the p53 gene are frequent in primary, resected non-small cell lung cancer. Oncogene 1990, **5**:1603-1610

Christensen PM, Gotzsche PC, Brosen K. The sparteine/debrisoquine (CYP2D6) oxidation polymorphism and the risk of lung cancer: a meta-analysis. Eur J Clin Pharmacol 1997, **51**:389-393

Chung KY, Mukhopadhyay T, Kim J, Casson A, Ro JY, Goepfert H, Hong WK, Roth JA. Discordant p53 gene mutations in primary head and neck cancers and corresponding second primary cancers of the upper aerodigestive tract. Cancer Res 1993, **53**:1676-1683

Chung GT, Sundaresan V, Hasleton P, Rudd R, Taylor R, Rabbitts PH. Sequential molecular genetic changes in lung cancer development. Oncogene 1995, **11**:2591-2598

Clayton F. The spectrum and significance of bronchioloalveolar carcinomas. Pathol Annu 1988, **23**:361-394

Cleary M, Smith S, Sklar J. Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin transcript resulting from the t(14:18) translocation. Cell 1986, **47**:19-28

Cline MJ, Battifora H. Abnormalities of protooncogenes in non-small cell lung cancer. Correlations with tumour type and clinical charcteristics. Cancer 1987, **60**:2669-2674

Coffer PJ, Jin J, Woodgett JR. Protein kinase B (c-Akt): a multifunctional mediator of phosphatidylinositol 3-kinase activation. Biochem J 1995, **335**:1-13

Cooper DA, Eldridge AL, Peters JC. Dietary carotenoids and lung cancer: a review of recent research. Nutr Rev 1999, **57**:133-145

Cormier RT, Hong KH, Halberg RB, Hawkins TL, Richardson P, Mulherkar R, Dove WF, Lander ES. Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis. Nat Genet 1997, **17**:88-91

Crawford HC, Fingleton BM, Rudolph-Owen LA, Goss KJ, Rubinfeld B, Polakis P, Matrisian LM. The metalloproteinase matrilysin is a target of beta-catenin transactivation in intestinal tumors. Oncogene 1999, **18**:2883-2891

Crechet JB, Bernardi A, Parmeggiani A. Distal switch II region of Ras2p is required for interaction with guanine nucleotide exchange factor. J Biol Chem 1996, **271**:17234-17240

Cripps KJ, Curtis LJ, Wyllie AH. Mutational analysis of the MCC gene by single-strand conformational polymorphism analysis. Eur J Cancer 1995, **31A**:853-855

Cross DA, Alessi DR, Cohen P, Andjelkovich M, Hemmings BA. Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. Nature 1995, **378**:785-789

Curtis LJ, Bubb VJ, Gledhill S. Morris RG, Bird CC, Wyllie AH. Loss of heterozygosity of MCC is not associated with mutation of the retained allele in sporadic colorectal cancer. Hum Mol Gen 1994, **3**:443-446

Daniel JM, Reynolds AB. Tyrosine phosphorylation and cadherin/catenin function. Bioessays 1997, **19**:883-891

Dartigues JF, Dabis F, Gros N, Moise A, Bois G, Salamon R, Dilhuydy JM, Courty G. Dietary vitamin A, beta carotene and risk of epidermoid lung cancer in south-western France. Eur J Epidemiol 1990, **6**:261-265

Derijard B, Hibi M, Wu IH, Barrett T, Su B, Deng T, Karin M, Davis RJ. JNK1: a protein kinase stimulated by UV light and Ha-Ras that binds and phosphorylates the c-Jun activation domain. Cell 1994, **76**:1025-1037

Dierick H, Bejsovec A. Cellular mechanisms of wingless/Wnt signal transduction. Curr Top Dev Biol 1999, **43**:153-190

Dietrich WF, Lander ES, Smith JS, Moser AR, Gould AG, Luongo C, Borenstein N, Dove W. Genetic identification of MOM-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. Cell 1993, **75**:631-639

Dipple A. DNA adducts of chemical carcinogens. Carcinogenesis 1995, **16**:437-441

Dobbie Z, Spycher M, Mary JL, Haner M, Guldenschuh I, Hurliman R, Amman R, Roth J, Muller H, Scott RJ. Correlation between the development of extracolonic manifestations in FAP patients and mutations beyond codon 1403 in the APC gene. J Med Genet 1996, **33**:274-280

Dobbie Z, Heinimann K, Bishop DT, Muller H, Scott RJ. Identification of a modifier gene locus on chromosome 1p35-36 in familial adenomatous polyposis. Hum Genet 1997, **99**:653-657

Doll R, Peto R. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong non-smokers. J Epidemiol Community Health 1978, **32**:303-313

Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. Natl Cancer Inst 1981, **66**:1191-1308

Dominguez I, Itoh K, Sokol SY. Role of glycogen synthase kinase 3 beta as a negative regulator of dorsoventral axis formation in Xenopus embryos. Proc Natl Acad Sci U S A 1995, **92**:8498-8502

Doran SE, Ren XD, Betz AL, Pagel MA, Neuwelt EA, Roessler BJ, Davidson BL. Gene expression from recombinant viral vectors in the central nervous system after blood-brain barrier disruption. Neurosurgery 1995, **36**:965-970

Dorsky RI, Moon RT, Raible DW. Control of neural crest cell fate by the Wnt signalling pathway. Nature 1998, **396**:370-373

Dosaka-Akita H, Katabami M, Hommura H, Fujioka Y, Katoh H, Kawakami Y. Bcl-2 expression in non-small cell lung cancers: higher frequency of expression in squamous cell carcinomas with earlier pT status. ncology 1999, **56**:259-264

Dove WF, Luongo C, Connelly CS, Gould KA, Shoemaker AR, Moser AR, Gardner RL. The adenomatous polyposis coli gene of the mouse in development and neoplasia. Cold Spring Harb Symp Quant Biol 1994;**59:**501-508

Doyle DA, Lee A, Lewis J, Kim E, Sheng M, MacKinnon R. Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. Cell 1996, **85**:1067-1076

Dubovsky J, Sheffield VC, Duyk GM, Weber JL. Sets of short tandem repeat polymorphisms for efficient linkage screening of the human genome. Hum Mol Genet 1995, **4**:449-452

Dunlop MG, Wyllie AH, Nakamura Y, Steel CM, Evans HJ, White RL, Bird CC. Genetic linkage map of six polymorphic DNA markers around the gene for familial adenomatous polyposis on chromosome 5. Am J Hum Genet 1990, **47**:982-987

Duro D, Bernard O, Della Valle V, Berger R, Larsen CJ. A new type of p16INK4/MTS1 gene transcript expressed in B-cell malignancies. Oncogene 1995, **11**:21-29

Dunsmore SE, Saarialho-Kere UK, Roby JD, Wilson CL, Matrisian LM, Welgus HG, Parks WC. Matrilysin expression and function in airway epithelium. J Clin Invest 1998, **102**:1321-1331

Dyson N. The regulation of E2F by pRB-family proteins. Genes Dev 1998, **12**:2245-2262

Ebinu JO, Bottorff DA, Chan EY, Stang SL, Dunn RJ, Stone JC. RasGRP, a Ras guanyl nucleotide- releasing protein with calcium- and diacylglycerol-binding motifs. Science 1998, **280**:1082-1086

Eccles DM, van der Luijt R, Breukel C, Bullman H, Bunyan D, Fisher A, Barber J, du Boulay C, Primrose J, Burn J, Fodde R. Hereditary desmoid disease due to a frameshift mutation at codon 1924 of the APC gene. **Am J Hum Genet** 1996, 59:1193-1201

Edwards CW. Pulmonary adenocarcinoma: Review of 106 cases and proposed new classification. J Clin Pathol 1987, 40:123-135

Erisman MD, Rothberg PG, Diehl RE, Morse CC, Spandorfer JM, Astrin SM. Deregulation of c-myc gene expression in human colon carcinoma is not accompanied by amplification or rearrangement of the gene. Mol Cell Biol 1985, **5**:1969-1976

van Ets JH, Kirkpatrick C, van de Wetering M, Molenaar M, Miles A, Kuipers J, Destree O, Peifer M, Clevers H. Identification of APC2, a homologue of the adenomatous polyposis coli tumour suppressor. Curr Biol 1999, **28**, 9:105-108

Evan GI, Littlewood TD. The role of c-myc in cell growth. Curr Opin Genet Dev 1993, **3**:44-49

Ewart JL, Cohen MF, Meyer RA, Huang GY, Wessels A, Gourdie RG, Chin AJ, Park SM, Lazatin BO, Villabon S, Lo CW. Heart and neural tube defects in transgenic mice overexpressing the Cx43 gap junction gene. Development 1997, **124**:1281-1292

Fagotto F, Funayama N, Gluck U, Gumbiner BM. Binding to cadherins antagonizes the signaling activity of beta-catenin during axis formation in Xenopus. J Cell Biol 1996, **132**:1105-1114

Fagotto F, Gluck U, Gumbiner BM. Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin. Curr Biol 1998, **8**:181-190

Fagotto F, Jho Eh, Zeng L, Kurth T, Joos T, Kaufmann C, Costantini F. Domains of axin involved in protein-protein interactions, Wnt pathway inhibition, and intracellular localization. J Cell Biol 1999, **145**:741-756

Farnsworth CL, Freshney NW, Rosen LB, Ghosh A, Greenberg ME, Feig LA. Calcium activation of Ras mediated by neuronal exchange factor Ras-GRF. Nature 1995, **376**:524-527

Fayad L, 'O' Brein S. Chronic lymphocytic leukemia and associated disorders, in Pazdur R (ed): Medical Oncology: A comprehensive Review (ed 2). New York, NY, PRR, Inc. 1995: 37-56

Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990, **61**:759-767

Feil R, Brocard J, Mascrez B, LeMeur M, Metzger D, Chambon P. Ligand-activated site-specific recombination in mice. Proc Natl Acad Sci *USA* 1996, **93**:10887-10890

Feinberg AP, Vogelstein B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 1983, **132**:6-13

Field JK, Spandidos DA. The role of ras and myc oncogenes in human solid tumours and their relevance in diagnosis and prognosis (review). Anticancer Res 1990, **10**:1-22

Fingleton BM, Heppner Goss KJ, Crawford HC, Matrisian LM. Matrilysin in early stage intestinal tumorigenesis. APMIS 1999, **107**:102-110

Fisher AL, Caudy M. Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates. Genes Dev 1998, **12**:1931-1940

Fodde R, Edelmann W, Yang K, van Leeuwen C, Carlson C, Renault B, Breukel C, Alt E, Lipkin M, Khan PM, et al. A targeted chain-termination mutation in the mouse Apc gene results in multiple intestinal tumors. Proc Natl Acad Sci *U S A* 1994, **91**:8969-8973

Fodde R, Khan PM. Genotype-phenotype correlations at the adenomatous polyposis coli (APC) gene. Crit Rev Oncog 1995,**6**:291-303

Foley KP, Eisenman RN. Two MAD tails: what the recent knockouts of Mad1 and Mxi1 tell us about the MYC/MAX/MAD network. Biochim Biophys Acta 1999, **1423**:M37-47

Fong KM, Zimmerman PV, Smith PJ. Correlation of loss of heterozygosity at 11p with tumour progression and survival in non-small cell lung cancer. Genes Chrom Cancer 1994, **10**:183-189

Fong KM, Zimmerman PV, Smith PJ. Tumour progression and loss of heterozygosity at 5q and 18q in non-small cell lung cancer. Cancer Res 1995a, **55**:220-223

Fong KM, Zimmerman PV, Smith PJ. Lung pathology. The molecular genetics of non-small cell lung cancer. Pathology 1995b, **27**:295-301

Franke WW, Goldschmidt MD, Zimbelmann R, Mueller HM, Schiller DL, Cowin P. Molecular cloning and amino acid sequence of human plakoglobin, the common junctional plaque protein. Proc Natl Acad Sci U S A 1989, **86**:4027-4031

Franke TF, Yang SI, Chan TO, Datta K, Kazlauskas A, Morrison DK, Kaplan DR, Tsichlis PN. The protein kinase encoded by the Akt proto-oncogene is a target of the PDGF-activated phosphatidylinositol 3-kinase. Cell 1995, **81**:727-736

Franklin WA, Gazdar AF, Haney J, Wistuba II, La Rosa FG, Kennedy T, Ritchey DM, Miller YE. Widely dispersed p53 mutation in respiratory epithelium. A novel mechanism for field carcinogenesis. J Clin Invest 1997, **100**:2133-2137

Fraser GE, Beeson WL, Phillips RL Diet and lung cancer in California Seventh-day Adventists. Am J Epidemiol 1991, **133**:683-693

Frayling IM, Beck NE, Ilyas M, Dove-Edwin I, Goodman P, Pack K, Bell J, Williams CB, Hodgson SV, Thomas HJ, Talbot IC, Bodmer WF, Tomlinson IP. The APC variants I130K and E1317Q are associated with colorectal tumours, but not always with a family history. Proc. Natl. Acad. Sci USA. 1998, **95**:10722-10727

Friberg L, Cederlof R. Late effects of air pollution with special reference to lung cancer. Environ Health Perspect 1978, **22**:45-66

Friedl W, Meuschel S, Caspari R, Lamberti C, Krieger S, Sengteller M, Propping P. Attenuated familial adenomatous polyposis due to a mutation in the 3' part of the APC gene. A clue for understanding the function of the APC protein. *Hum Genet* 1996, **97**:579-584

Fujita-Yoshigaki J, Shirouzu M, Ito Y, Hattori S, Furuyama S, Nishimura S, Yokoyama S. A constitutive effector region on the C-terminal side of switch I of the Ras protein. J Biol Chem 1995, **270**:4661-4667

Fukasawa K, Taesaeng C, Kuriyama R, Rulong S, Vande Woude GF. Abnormal centrosome amplification in the absence of p53. Science 1996, **271**:1744-1747

Funayama N, Fagotto F, McCrea P, Gumbiner BM. Embryonic axis induction by the armadillo repeat domain of beta-catenin: evidence for intracellular signaling. J Cell Biol 1995, **128**:959-968

Furth ME, Aldrich TH, Cordon-Cardo C. Expression of ras proto-oncogene proteins in normal human tissues. Oncogene 1987, **1**:47-58

Gaidano G, Ballerini P, Gong JZ, Inghirami G, Neri A, Newcomb EW, Magrath IT, Knowles DM, Dalla-Favera R p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 1991, **88**:5413-5417

Galceran J, Farinas I, Depew MJ, Clevers H, Grosschedl R. Wnt3a-/--like phenotype and limb deficiency in Lef1(-/-)Tcf1(-/-) mice. Genes Dev 1999, **13**:709-717

Gao HG, Chen JK, Stewart J, Song B, Rayappa C, Whong WZ, Ong T. Distribution of p53 and K-ras mutations in human lung cancer tissues. Carcinogenesis 1997, **18**:473-478

Gastier JM, Pulido JC, Sunden S, Brody T, Buetow KH, Murray JC, Weber JL, Hudson TJ, Sheffield VC, Duyk GM. Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. Hum Mol Genet 1995, **4**:1829-1836

Gavin BJ, McMahon JA, McMahon AP. Expression of multiple novel Wnt-1/int-1-related genes during fetal and adult mouse development. Genes Dev 1990, **4**:2319-2332

Gavin BJ, McMahon AP. Differential regulation of the Wnt gene family during pregnancy and lactation suggests a role in postnatal development of the mammary gland. Mol Cell Biol 1992, **12**:2418-2423

Gazdar AF, McDowell . Pathobiology of lung cancers. In: biology of lung cancer; diagnosis and treatment 1988

Gazdar AF, Carbone DP. The biology and molecular genetics of lung cancer. 1994 RG Landes, Austin,Texas.

Gazzeri S, Gouyer V, Vour'ch C, Brambilla C, Brambilla E. Mechanisms of p16INK4A inactivation in non-small-cell lung cancers. Oncogene 1998, **16**:497-504

Gealy R, Zhang L, Siegfried JM, Luketich JD, Keohavong P. Comparison of mutations in the p53 and K-ras genes in lung carcinomas from smoking and nonsmoking women. Cancer Epidemiol Biomarkers Prev 1999, **8**:297-302

van Genderen C, Okamura RM, Farinas I, Quo RG, Parslow TG, Bruhn L, Grosschedl R Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice. Genes Dev 1994, **8**:2691-2703

Giese K, Kingsley C, Kirshner JR, Grosschedl R. Assembly and function of a TCR alpha enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. Genes Dev 1995, **9**:995-1008

Giglione C, Parmeggiani A. Raf-1 is involved in the regulation of the interaction between guanine nucleotide exchange factor and Ha-ras. Evidences for a function of Raf-1 and phosphatidylinositol 3-kinase upstream to Ras. J Biol Chem 1998, **273**:34737-3444

Glavac D, Dean M. Optimization of the single-strand conformation polymorphism (SSCP) technique for detection of point mutations. Hum Mutat 1993, **2**:404-414

Gleeson CM, Sloan JM, McGuigan JA, Ritchie AJ, Weber JL, Russell SE. Barrett's oesophagus: microsatellite analysis provides evidence to support the proposed metaplasia-dysplasia-carcinoma sequence. Genes Chromosomes Cancer 1998, **21**:49-60

Gonzalez FJ, Vilbois F, Hardwick JP, McBride OW, Nebert DW, Gelboin HV, Meyer UA. Human debrisoquine 4-hydroxylase (P45011D1): cDNA and deduced amino acid sequence and assignment of the CYP2D locus to chromosome 22. Genomics 1988, **2**:174-179

Gonzalez MV, Artimez ML, Rodrigo L, Lopez-Larrea C, Menendez MJ, Alvarez V, Perez R, Fresno MF, Perez MJ, Sampedro A, Coto E. Mutation analysis of the p53, APC, and p16 genes in the Barrett's oesophagus, dysplasia, and adenocarcinoma. J Clin Pathol 1997, **50**:212-217

Goodenough DA, Goliger JA, Paul DL. Connexins, connexons, and intercellular communication. Annu Rev Biochem 1996, **65:**475-502

Gorgoulis VG, Zacharatos P, Kotsinas A, Liloglou T, Kyroudi A, Veslemes M, Rassidakis A, Halazonetis T, Field JK, Kittas C. Alterations of thep16-pRB pathway and the chromosome locus 9p21-22 in non-small cell lung carcinomas. Relationship with p53 and MDM2 protein expression. Am J Pathol 1998, **153**:1749-1765

Goyette MC, Cho K, Fasching CL, Levy DB, Kinzler KW, Paraskeva C, Vogelstein B, Stanbridge EJ. Progression of colorectal cancer is associated with multiple tumor suppressor gene defects but inhibition of tumorigenicity is accomplished by correction of any single defect via chromosome transfer. Mol Cell Biol 1992, **12**:1387-1395

Grandori C, Eisenman RN. Myc target genes. Trends Biochem Sci 1997, **22**:177-181

Graziano SL, Gamble GP, Newman NB, Abbott LZ, Rooney M, Mookherjee S, Lamb ML, Kohman LJ, Poiesz BJ. Prognostic significance of K-ras codon 12 mutations in patients with resected stage I and II non-small-cell lung cancer. J Clin Oncol 1999, **17**:668-675

Greenblatt MS, Bennett WP, Hollstein M, Harris CC. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res 1994, **54**:4855-4878

Groden J, Thilveris A, Samowitz W, Carlson M, Gelbert L, Albertsen H, Joslyn G, Stevens J, Spiro L, Robertson M, Sargeant L, Krapcho K, Wolff E, Burt R, Hughes JP, Warrington J, McPherson J, Wasmuth J, Le Paslier D, Abderrahlm H, Cohen D, Leppert M, White R. Identification and characterization of the familial adenomatous polyposis coli gene. Cell 1991, **66**:589-600

Groden J, Joslyn G, Wade S, Jones D, Bhattacharyya N, Spiro L, Thlivers A, Robertson M, Egan S, Meuth M, White R. Response of colon cancer cell lines to the introduction of APC, a colon-specific tumour suppressor gene. Cancer Res 1995, **55**:1531-1539

Guengerich FP. Comparisons of catalytic selectivity of cytochrome P450 subfamily enzymes from different species. Chem Biol Interact 1997, **106**:161-182

Gumbiner BM. Proteins associated with the cytoplasmic surface of adhesion molecules. Neuron 1993, **11**:551-564

Gumbiner BM. Cell adhesion: the molecular basis of tissue architecture and morphogenesis. Cell 1996, **84**:345-357

Haegel H, Larue L, Ohsugi M, Fedorov L, Herrenknecht K, Kemler R. Lack of beta-catenin affects mouse development at gastrulation. Development 1995, **121**:3529-3537

Halazonetis TD, Kandil AN. Determination of the c-MYC DNA-binding site. Proc Natl Acad Sci U S A 1991, **88**:6162-6166

Hall M, Peters G. Genetic alterations of cyclins, cyclin-dependent kinases, and Cdk inhibitors in human cancer. Adv Cancer Res 1996, **68**:67-108

Hamada F, Tomoyasu Y, Takatsu Y, Nakamura M, Nagai S, Suzuki A, Fujita F, Shibuya H, Toyoshima K, Ueno N, Akiyama T. Negative regulation of wingless signaling by D-axin, a drosophila homolog of axin. Science 1999, **283:**1739-1742

Hamilton M, Wolfman A. Ha-ras and N-ras regulate MAPK activity by distinct mechanisms in vivo. Oncogene 1998, **16**:1417-1428

Hannon GJ, Beach D. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. Nature 1994, **371**:257-261

Harber DA. Splicing into scenecence:the curious case of p16 and p19ARF. Cell, 1997, **91**:551-558

Hart MJ, de los Santos R, Albert IN, Rubinfeld B, Polakis P. Downregulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. Curr Biol 1998, **8**:573-81

Haskins J, Gu L, Wittchen ES, Hibbard J, Stevenson BR. ZO-3, a novel member of the MAGUK protein family found at the tight junction, interacts with ZO-1 and occludin. J Cell Biol 1998, **141**:199-208

Harvey JJ. An unidentified virus whivh causes the rapid production of tumours in mice. Nature, 1964, **204**:1104-1105

Hatzfeld M. The armadillo family of structural proteins. Int Rev Cytol 1999, **186**:179-224

Havre PA, Yuan J, Hedrick L, Cho KR, Glazer PM. p53 inactivation by HPV16 E6 results in increased mutagenesis in human cells. Cancer Res 1995, **55**:4420-4424

Hayashi K. PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. PCR Methods Appl 1991, **1**:34-38

Hayashi K, Yandell DW. How sensitive is PCR-SSCP? Hum Mutat 1993, **2**:338-346

Hayashi N, Sugimoto Y, Tsuchiya E, Ogawa M, Nakamura Y. Somatic mutation of the MTS (multiple tumour suppressor) gene 1/CDKN4I cyclin dependent kinase 4 inhibitor) gene in human primary non-small cell lung carcinomas Biochem Biophys Res Comm 1994, **202**:1426-1436

Hayashi S, Rubinfeld B, Souza B, Polakis P, Wieschaus E, Levine AJ. A Drosophila homolog of the tumor suppressor gene adenomatous polyposis coli down-regulates beta-catenin but its zygotic expression is not essential for the regulation of Armadillo. Proc Natl Acad Sci *USA* 1997, **94**:242-247

He X, Saint-Jeannet JP, Woodgett JR, Varmus HE, Dawid IB. Glycogen synthase kinase-3 and dorsoventral patterning in Xenopus embryos. Nature 1995, **374**:617-622

He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT, Morin PJ, Vogelstein B, Kinzler KW. Identification of c-MYC as a target of the APC pathway. Science 1998, **281**:1509-1512

Hecht SS, Hoffmann D. Tobacco-specific nitrosamines, an important group of carcinogens in tobacco and tobacco smoke. Carcinogenesis 1988, **9**:875-884

Hecht SS. Tobacco smoke carcinogens and lung cancer. J Natl Cancer Inst 1999, **91**:1194-1210

Heighway J, Thatcher N, Cerny T, Hasleton PS. Genetic predisposition to human lung cancer. Br J Cancer 1986, **53**:453-457

Heighway J, Hoban PR, Wyllie AH. Ssp1 polymorphism in sequence encoding 3' untranslated region of the APC gene. Nucleic Acids Res 1991, **19**:6966

Heighway J. Hoban PR, Wyllie AH. Mae111 polymorphism in sequence encoding 3' untranslated region of the MCC gene. Nucleic Acids Res 1992, **20**:1160

Henriksson M, Luscher B. Proteins of the Myc network: essential regulators of cell growth and differentiation. Adv Cancer Res 1996, **68**:109-182

Hermiston ML, Gordon JI. In vivo analysis of cadherin function in the mouse intestinal epithelium: essential roles in adhesion, maintenance of differentiation, and regulation of programmed cell death. J Cell Biol 1995, **129**:489-506

Hermiston ML, Wong MH, Gordon JI. Forced expression of E-cadherin in the mouse intestinal epithelium slows cell migration and provides evidence for nonautonomous regulation of cell fate in a self-renewing system. Genes Dev 1996, **10**:985-996

Herskowitz I. Functional inactivation of genes by dominant negative mutations. Nature 1987, **329**:219-222

Herwig S, Strauss M. The retinoblastoma protein: a master regulator of cell cycle, differentiation and apoptosis. Eur J Biochem 1997, **246**:581-601

Herzlinger D, Qiao J, Cohen D, Ramakrishna N, Brown AM. Induction of kidney epithelial morphogenesis by cells expressing Wnt-1. Dev Biol 1994, **166**:815-818

van der Heyden MA, Rook MB, Hermans MM, Rijksen G, Boonstra J, Defize LH, Destree OH. Identification of connexin43 as a functional target for Wnt signalling. Cell Sci 1998, **111**:1741-1749

Higashiyama M, Doi O, Kodama K,Yokouchi H, Nakamori S, Tateishi R. Bcl-2 oncoprotein in surgically resected nonsmall cell lung cancer: possibly favorable prognostic factors in association with low incidence of distant metastasis. J Surg Oncol 1997, **64**:48-54

Hiltunen MO, Alhonen L, Koistinaho J, Myohanen S, Paakkonen M, Marin S, Kosma VM, Janne J. Hypermethylation of the APC (adenomatous polyposis coli) gene promoter region in human colorectal carcinoma. Int J Cancer 1997, **70**:644-648

Hinck L, Nathke IS, Papkoff J, Nelson WJ. Dynamics of cadherin/catenin complex formation: novel protein interactions and pathways of complex assembly. J Cell Biol 1994, **125**:1327-1340

Hirano T, Franzen B, Kato H, Ebihara Y, Auer G. Genesis of squamous cell lung carcinoma. Sequential changes of proliferation, DNA ploidy, and p53 expression. Am J Pathol 1994, **144**:296-302

Hochman L, Segev N, Sternberg N, Cohen G. Site-specific recombinational circularization of bacteriophage P1 DNA. Virology 1983, **131**:11-17

Hochstrasser M. Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. Curr Opin Cell Biol 1995, 7:215-223

Hoess RH, Abremski K. Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. J Mol Biol 1985, **181**:351-362

Hoess RH, Wierzbicki A, Abremski K. The role of the loxP spacer region in P1 site-specific recombination. Nucleic Acids Res 1986, **14**:2287-2300

Hoffmann D, Adams JD, Haley NJ. Reported cigarette smoke values: a closer look. Am J Public Health 1983, **73**:1050-1053

Hoffmann D, Hoffmann I. The changing cigarette, 1950-1995. J Toxicol Environ Health 1997, **50**:307-364

Hoffman B, Liebermann DA. The proto-oncogene c-myc and apoptosis. Oncogene 1998, **17**:3351-3357

Hogan B, Beddington R, Constantini F, Lacey E. Manipulating the Mouse Embryo: A Laboratory Manual (second addition) 1994, Cold Spring Harbor Laboratory Press.

Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 mutations in human cancers. Science 1991, **253**:49-53

Horii A, Nakatsuru S, Miyoshi Y, Ichii S, Nagase H, Ando H, Yanagisawa A, Tsuchiya E, Kato Y, Nakamura Y. Frequent somatic mutations of APC gene in human pancreatic cancer. Cancer Res 1992a, **52**:6696-6698

Horii A, Nakatsuru S, Miyoshi Y, Ichii S, Nagase H, Kato Y, Yanagisawa A, Nakamura Y. The APC gene, responsible for familial adenomatous polyposis, is mutated in human gastric cancer. Cancer Res 1992b, **52**:3231-3233

Horii A, Nakatsuru S, Ichii S, Nagase H, Nakamura Y. Multiple forms of the APC gene transcripts and their tissue-specific expression. Hum Mol Genet 1993, **2**:283-287

Hosoe S, Ueno K, Shigedo Y, Tachibana I, Osaki T, Kumagai T, Tanio Y, Kawase I, Nakamura Y, Kishimoto T. A frequent deletion of chromosome 5q21 in advanced small cell and non small cell carcinoma of the lung. Cancer Res 1994, **54**:1787-1790

Houldsworth J, Chaganti RS. Comparative genomic hybridization: an overview. Am J Pathol 1994, **145**:1253-1260

Hsu SC, Galceran J, Grosschedl R. Modulation of transcriptional regulation by LEF-1 in response to Wnt-1 signaling and association with beta-catenin. Mol Cell Biol 1998, **18**:4807-4818

Hsu W, Zeng L, Costantini F. Identification of a domain of Axin that binds to the serine/threonine protein phosphatase 2A and a self-binding domain. J Biol Chem 1999, **274**:3439-3445

Huang CL, Taki T, Adachi M, Konishi T, Higashiyama M, Kinoshita M, Hadama T, Miyake M. Mutations of p53 and K-ras genes as prognostic factors for non-small cell lung cancer. Int J Oncol 1998, **12**:553-563

Huber O, Korn R, McLaughlin J, Ohsugi M, Herrmann BG, Kemler R. Nuclear localization of beta-catenin by interaction with transcription factor LEF-1. Mech Dev 1996, **59**:3-10

Hulsken J, Behrens J, Birchmeier W. Tumor-suppressor gene products in cell contacts: the cadherin-APC-armadillo connection. Curr Opin Cell Biol 1994a, **6**:711-716

Hulsken J, Birchmeier W, Behrens J. E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton. J Cell Biol 1994b, **127**:2061-2069

Hung J, Kishimoto Y, Sugio K, Virmani A, McIntire DD, Minna JD, Gazdar AF. Allele-specific chromosome 3p deletions occur at an early stage in the pathogenesis of lung carcinoma. JAMA 1995, **273**:1908

Hunt CA, Schenker LJ, Kennedy MB. PSD-95 is associated with the postsynaptic density and not with the presynaptic membrane at forebrain synapses. J Neurosci 1996, **16**:1380-1388

Hunter T, Pines J. Cyclins and cancer. II: Cyclin D and CDK inhibitors come of age. Cell 1994, **79**:573-582

Hurlin PJ, Queva C, Eisenman RN. Mnt, a novel Max-interacting protein is coexpressed with Myc in proliferating cells and mediates repression at Myc binding sites. Genes Dev 1997, **11**:44-58

Hussussian CJ, Struewing JP, Goldstein AM, Higgins PA, Ally DS, Sheahan MD, Clark WH Jr, Tucker MA, Dracopoli NC. Germline p16 mutations in familial melanoma. Nat Genet 1994, **8**:15-21

Hyman AA, Karsenti E. Morphogenetic properties of microtubules and mitotic spindle assembly. Cell 1996, **84**:401-410

Ichii S, Horii A, Nakatsuru S, Furuyama J, Utsunomiya J, Nakamura Y. Inactivation of both APC alleles in an early stage of colon adenomas in a patient with familial adenomatous polyposis (FAP). Hum Mol Genet 1992, **1**:387-390

Ikeda MA, Jakoi L, Nevins JR. A unique role for the Rb protein in controlling E2F accumulation during cell growth and differentiation. Proc Natl Acad Sci U S A 1996, **93**:3215-3220

Ikeda S, Kishida S, Yamamoto H, Murai H, Koyama S, Kikuchi A. Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3beta and beta-catenin and promotes GSK-3beta-dependent phosphorylation of beta-catenin. EMBO J 1998, **17**:1371-1384

Imai Y, Oda H, Tsurutani N, Nakatsuru Y, Inoue T, Ishikawa T. Frequent somatic mutations of the APC and p53 genes in sporadic ampullary carcinomas. Jpn J Cancer Res 1997, **88**:846-854

Imhof A, Yang XJ, Ogryzko VV, Nakatani Y, Wolffe AP, Ge H. Acetylation of general transcription factors by histone acetyltransferases. Curr Biol 1997, 7:689-692

Imlay JA, Linn S. DNA damage and oxygen radical toxicity. Science 1988, **240**:1302-1309

Ingham PW. Has the quest for a Wnt receptor finally frizzled out? Trends Genet 1996, **12**:382-384

Inoue S, Salmon ED. Force generation by microtubule assembly/disassembly in mitosis and related movements. Mol Biol Cell 1995, **6**:1619-1640

Irminger-Finger I, Laymon RA, Goldstein LS. Analysis of the primary sequence and microtubule-binding region of the Drosophila 205K MAP. J Cell Biol 1990, **111**:2563-2572

Itoh M, Nagafuchi A, Moroi S, Tsukita S. Involvement of ZO-1 in cadherin-based cell adhesion through its direct binding to alpha catenin and actin filaments. J Cell Biol 1997, **138**:181-192

Itoh M, Morita K, Tsukita S. Characterization of ZO-2 as a MAGUK family member associated with tight as well as adherens junctions with a binding affinity to occludin and alpha catenin. J Biol Chem 1999, **274**:5981-5986

Jacobs-Cohen RJ, Spiegelman M, Cookingham JC, Bennett D. Knobbly, a new dominant mutation in the mouse that affects embryonic ectoderm organization. Genet Res 1984, **43**:43-50

Jen J, Powell SM, Papadopoulos N, Smith KJ, Hamilton SR, Vogelstein B, Kinzler KW. Molecular determinants of dysplasia in colorectal lesions. Cancer Res 1994, **54**:5523-5526

Jiang W, Kahn SM, Guillem JG, Lu SH, Weinstein IB. Rapid detection of ras oncogenes in human tumors: applications to colon, esophageal, and gastric cancer. Oncogene 1989, **4**:923-928

Jiang SX, Kameya T, Sato Y, Yanase N, Yoshimura H, Kodama T. Bcl-2 protein expression in lung cancer and close correlation with neuroendocrine differentiation. Am J Pathol 1996, **148**:837-846

Johnson L, Greenbaum D, Cichowski K, Mercer K, Murphy E, Schmitt E, Bronson RT, Umanoff H, Edelmann W, Kucherlapati R, Jacks T. K-ras is an essential gene in the mouse with partial functional overlap with N-ras. Genes Dev 1997, **11**:2468-2481

Jones MK, Jackson JH. Ras-GRF activates Ha-Ras, but not N-Ras or K-Ras 4B, protein in vivo. Biol Chem 1998, **273**:1782-1787

Joslyn G, Calson M, Thliveris A, Halbertsen A, Gelbert L, Samowitz W, Groden J, Stevens J, Spiro L, Robertson M, Sargeant L, Krapcho K, Wolff E, Burt R, Hughes JP, Warrington J, McPherson J, Wasmuth J, LePaslier D, Abderrahim H, Cohen D, Leppert M, White R. Identification of deletion mutations and three new genes at the familial polyposis locus. Cell 1991, **66**:601-613

Joslyn G, Richardson DS, White R, Alber T. Dimer formation by an N-terminal coiled coil in the APC protein. Proc Natl Acad Sci U S A 1993, **90**:11109-11113

Juwana JP, Henderikx P, Mischo A, Wadle A, Fadle N, Gerlach K, Arends JW, Hoogenboom H, Pfreundschuh M, Renner C. EB/RP gene family encodes tubulin binding proteins. Int J Cancer 1999, **81**:275-284

Kadonaga JT. Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. Cell 1998, **92**:307-313

Kamb A, Gruis NA, Weaver-Feldhaus J, Liu Q, Harshman K, Tavtigan SV, Stockert E, Day RSIII, Johnson BE, Skolnick MH. A cell cycle regulator potentially involved in genesis of many tumour types. Science 1994, **264**:436-440

Kamijo T, Zindy F, Roussel MF, Quelle DE, Downing JR, Ashmun RA, Grosveld G, Sherr CJ. Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. Cell 1997, **91**:649-659

Kanegae Y, Lee G, Sato Y, Tanaka M, Nakai M, Sakaki T, Sugano S, Saito I. Efficient gene activation in mammalian cells by using recombinant adenovirus expressing site-specific Cre recombinase. Nucleic Acids Res 1995, **23**:3816-3821

Kapeller R, Cantley LC. Phosphatidylinositol 3-kinase. Bioessays 1994, **16**:565-576

Kasai H, Nishimura S. Hydroxylation of guanine in nucleosides and DNA at the C-8 position by heated glucose and oxygen radical-forming agents. Environ Health Perspect 1986, **67**:111-116

Kashiwaba M, Tamura G, Ishida M. Aberrations of the APC gene in primary breast carcinoma. J Cancer Res Clin Oncol 1994, **120**:727-731

Kashiwabara K, Oyama T, Sano T, Fukuda T, Nakajima T. Correlation between methylation status of the p16/CDKN2 gene and the expression of p16 and Rb proteins in primary non-small cell lung cancers. Int J Cancer 1998, **79**:215-220

Kass-Eisler A, Falck-Pedersen E, Elfenbein DH, Alvira M, Buttrick PM, Leinwand LA. The impact of developmental stage, route of administration and the immune system on adenovirus-mediated gene transfer. Gene Ther 1994, **1**:395-402

Katz ME, McCormick F. Signal transduction from multiple Ras effectors. Curr Opin Genet Dev 1997, **7**:75-79

Kellendonk C, Tronche F, Monaghan AP, Angrand PO, Stewart F, Schutz G. Regulation of Cre recombinase activity by the synthetic steroid RU 486. Nucleic Acids Res 1996, **24**:1404-1411

Kelly K, Siebenlist U. Immediate-early genes induced by antigen receptor stimulation. Curr Opin Immunol 1995, **7**:327-332

Kelsey KT, Spitz MR, Zuo ZF, Wiencke JK. Polymorphisms in the glutathione S-transferase class mu and theta genes interact and increase susceptibility to lung cancer in minority populations (Texas United States). Cancer Causes Control 1997, **8**:554-559

Kemler R. Classical cadherins. Semin Cell Biol 1992, **3**:149-155

Kengaku M, Capdevila J, Rodriguez-Esteban C, De La Pena J, Johnson RL, Belmonte JCI, Tabin CJ. Distinct WNT pathways regulating AER formation and dorsoventral polarity in the chick limb bud. Science 1998, **280**:1274-1277

Kennedy MB. The postsynaptic density.. Curr Opin Neurobiol 1993, **3**:732-737

Keohavong P, DeMichele MA, Melacrinos AC, Landreneau RJ, Weyant RJ, Siegfried JM. Detection of K-ras mutations in lung carcinomas: relationship to prognosis. Clin Cancer Res 1996, **2**:411-418

Kern JA, Schwartz DA, Nordberg JE, Weiner DB, Greene MI, Torney L, Robinson RA. p185neu expression in human lung adenocarcinomas predicts shortened survival. Cancer Res 1990, **50**:5184-5187

Kern JA, Robinson RA, Gazdar A, Torney L, Weiner DB. Mechanisms of pl85HER2 expression in human non-small cell lung cancer cell lines. Am J Resir Cell Mol Biol 1992, **6**:359-363

Kerr K, Carey F, King G, Lamb D. Atypical alveolar hyperplasia: Relationship with pulmonary adenocarcinoma, p53, and c-erb B-2 expression. J Path 1994, **174**:249-256

Kim SH, Roth KA, Moser AR, Gordon JI. Transgenic mouse models that explore the multistep hypothesis of intestinal neoplasia. J Cell Biol 1993, **123**:877-893

Kim SO, Hasham MI, Katz S, Pelech SL. Insulin-regulated protein kinases during postnatal development of rat heart. J Cell Biochem 1998, **71**:328-339

Kimelberg HK, Norenberg MD. Astrocytes. Sci Am 1989, **260**:66-72

Kimmelman A, Tolkacheva T, Lorenzi MV, Osada M, Chan AM. Identification and characterization of R-ras3: a novel member of the RAS gene family with a non-ubiquitous pattern of tissue distribution. Oncogene 1997, **15**:2675-2685

Kinoshita I, Dosaka-Akita H, Mishina T, Akie K, Nishi M, Hiroumi H, Hommura F, Kawakami Y. Altered p16INK4A and retinoblastoma protein status in non-small cell lung cancer: Potential synergistic effect with altered p53 protein on proliferative activity. Cancer Res 1996, **56**:5557-5562

Kinzler KW, Nilbert MC, Volgelstein B, Bryon TM, Levy DB, Smith KJ, Preisinger AC, Hamilton SR, Hedge P, Markham A, Carlson M, Joslyn G, Groden J, White R, Miki Y, Miyoshi Y, Nishisho I, Nakamura Y. Identification of a gene located at chromosome 5q21 that is mutated in colorectal cancers. Science 1991a, **251**:1366-1370

Kinzler KW, Nilbert MC, Su L, Volgelstein B, Bryon T, Levy M, Smith KJ, Preisinger AC, Hedge P, McKechnie D, Finniear R, Markham A, Groffen J, Boguski MS, Altschul SF, Horri A, Ando H, Miyoshi Y, Miki Y, Nishisho I, Nakamura Y. Identification of FAP locus genes from chromosome 5q21. Science 1991b, **253**:661-664

Kinzler KW, Vogelstein B. Life (and death) in a malignant tumour. Nature 1996, **379**:19-20

Kirsch DG, Kastan MB. Tumor-suppressor p53: implications for tumor development and prognosis. J Clin Oncol 1998, **16**:3158-3168

Kirsten WH, Mayer LA. Morphological responses to a murine erythroblastosis virus. J Natl Cancer Inst 1967, **39**:311-335

Kishida S, Yamamoto H, Ikeda S, Kishida M, Sakamoto I, Koyama S, Kikuchi A. Axin, a negative regulator of the wnt signaling pathway, directly interacts with adenomatous polyposis coli and regulates the stabilization of beta-catenin. J Biol Chem 1998, **273**:10823-10826

Kishida M, Koyama S, Kishida S, Matsubara K, Nakashima S, Higano K, Takada R, Takada S, Kikuchi A Axin prevents Wnt-3a-induced accumulation of beta-catenin. Oncogene 1999, **18**:979-985

Kishimoto Y, Sugio K, Hung JY, Virmani AK, McIntire DD, Minna JD, Gazdar AF. Allele-specific loss in chromosome 9p loci in preneoplastic lesions accompanying non-small-cell lung cancers. J Natl Cancer Inst 1995, **87**:1224-1229

Kitamura H, Kameda Y, Nakamura N, Inayama Y, Nakatani Y, Shibagaki T, Ito N, Hayashi H, Kimura H, Kanisawa M. Atypical adenomatous hyperplasia an bronchoalveolar lung carcinoma. Analysis by morphometry and the expression of p53 and carcinoembryonic antigen. Am J Surg Pathol 1996, **20**:553-562

Klingensmith J, Yang Y, Axelrod JD, Beier DR, Perrimon N, Sussman DJ. Conservation of dishevelled structure and function between flies and mice: isolation and characterization of Dvl2. Mech Dev 1996, **58**:15-26

Knall C, Johnson GL. G-protein regulatory pathways: rocketing into the twenty-first century. J Cell Biochem Suppl 1998, **30-31**:137-46

Knekt P, Jarvinen R, Seppanen R, Rissanen A, Aromaa A, Heinonen OP, Albanes D, Heinonen M, Pukkala E, Teppo L. Dietary antioxidants and the risk of lung cancer. Am J Epidemiol 1991, **134**:471-479

Knudsen KA, Soler AP, Johnson KR, Wheelock MJ. Interaction of alpha-actinin with the cadherin/catenin cell-cell adhesion complex via alpha-catenin. J Cell Biol 1995, **130**:67-77

Knudson AG Jr. Retinoblastoma: A prototypic hereditary neoplasm. Semin Oncol 1978, **5**:57-60

Kodama T, Biyajima S, Watanabe S, Shimosato Y. Morphometric study of adenocarcinomas and hyperplastic epithelial lesions in the peripheral lung. Am J Clin Pathol 1986, **85**:146-151

Kohno H, Hiroshima K, Toyozaki T, Fujisawa T, Ohwada H. p53 mutation and allelic loss of chromosome 3p, 9p of preneoplastic lesions in patients with nonsmall cell lung carcinoma. Cancer 1999, **85**:341-347

Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, Kinzler KW, Vogelstein B, Clevers H. Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. Science 1997, **275**:1784-1787

Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ, Clevers H. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. Nat Genet 1998, **19**:379-383

Kozak M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res 1984, **12**:857-872

Kratzke RA, Greathens TM, Rubins JB, Maddaus MA, Niewoehner DE, Niehans GA, Geradts J. Rb and p16INK4a expression in resected non-small cell lung tumours. Cancer Res 1996, **56**:3415-3420

Kreyberg L. Lung cancer in workers in a nickel refinery. *Br J Ind Med* 1978 May;35(2):109-16

Kuhl M, Wedlich D. Wnt signalling goes nuclear. Bioessays 1997, **19**:101-104

Kung ITM, Lui IOL, Loke SL, Khim MA, Mok CK, Lamky SO SY. Pulmonary scar cancer: a pathologic reappraisal. Am J Surg Pathol 1985, **9**:391-400

Kurasono Y, Ito T, Kameda Y, Nakamura N, Kitamura H. Expression of cyclin D1, retinoblastoma gene protein, and p16 MTS1 protein in atypical adenomatous hyperplasia and adenocarcinoma of the lung. An immunohistochemical analysis. Virchows Arch 1998, **432**:207-215

Laken SJ, Petersen GM, Gruber SB, Oddoux C, Ostrer H, Giardiello FM, Hamilton SR, Hampel H, Markowitz A, Klimstra D, Jhanwar S, Winawer S, Offit K, Luce MC, Kinzler KW, Vogelstein B. Familial colorectal cancer in Ashenazim due to a hypermutable tract in APC. Nature Genetics 1997, **17**:79-83

Lane D. p53, guardian of the genome. Nature 1992, **358**:15-16

Lane J, Allan V. Microtubule-based membrane movement. Biochim Biophys Acta 1998, **1376**:27-55

Larue L, Ohsugi M, Hirchenhain J, Kemler R. E-cadherin null mutant embryos fail to form a trophectoderm epithelium. Proc Natl Acad Sci *USA* 1994, **91**:8263-8267

Lakso M, Sauer B, Mosinger B Jr, Lee EJ, Manning RW, Yu SH, Mulder KL, Westphal H. Targeted oncogene activation by site-specific recombination in transgenic mice. Proc Natl Acad Sci *USA* 1992 , **89**:6232-6236

Lakso M, Pichel JG, Gorman JR, Sauer B, Okamoto Y, Lee E, Alt FW, Westphal H. Efficient in vivo manipulation of mouse genomic sequences at the zygote stage. Proc Natl Acad Sci *USA* 1996, **93**:5860-5865

Laudanski J, Chyczewski L, Niklinska WE, Kretowska M, Furman M, Sawicki B, Niklinski J. Expression of bcl-2 protein in non-small cell lung cancer: correlation with clinicopathology and patient survival. Neoplasma 1999, **46**:25-30

Laurent-Puig P, Beroud C, Soussi T APC gene: database of germline and somatic mutations in human tumors and cell lines. Nucleic Acids Res 1998, **26**:269-270

Lavia P, Jansen-Durr P. E2F target genes and cell-cycle checkpoint control. Bioessays 1999, **21**:221-230

Lavieille JP, Gazzeri S, Riva C, Reyt E, Brambilla C, Brambilla E. p53 mutations and p53, Waf-1, Bax and Bcl-2 expression in field cancerization of the head and neck. Anticancer Res 1998, **18**:4741-4749

Lecuit T, Cohen SM. Proximal-distal axis formation in the Drosophila leg. Nature 1997, **388**:139-145

Lee AM, Fraumeni JF Jr. Arsenic and respiratory cancer in man: an occupational study. J Natl Cancer Inst 1969, **42**:1045-1052

Lee WH, Brookstein R, Hong F, Young LJ, Shew JY, Lee EY. Human retinoblastoma susceptibility gene:Cloning, identification, and sequence. Science 1987, **235**:1394-1399

Lee G. Non-motor microtubule-associated proteins. Curr Opin Cell Biol 1993, **5**:88-94

Legrand M, Stucker I, Marez D, Sabbagh N, LoGuidice JM, Broly F. Influence of a mutation reducing the catalytic activity of the cytochrome P450 CYP2D6 on lung cancer susceptibility. Carcinogenesis 1996, **17**:2267-2269

Lemaire P, Garrett N, Gurdon JB. Expression cloning of Siamois, a Xenopus homeobox gene expressed in dorsal-vegetal cells of blastulae and able to induce a complete secondary axis. Cell 1995, **81**:85-94

Leon J, Guerrero I, Pellicer A. Differential expression of the ras gene family in mice. Mol Cell Biol 1987, **7**:1535-1540

Leppert M, Dobbs M, Scambler P, O'Connell P, Nakamura Y, Stauffer D, Woodward S, Burt R, Hughes J, Gardner E, et al The gene for familial polyposis coli maps to the long arm of chromosome 5. Science 1987, **238**:1411-1413

Leppert M, Burt R, Hughes JP, Samowitz W, Nakamura Y, Woodward S, Gardner E, Lalouel JM, White R. Genetic analysis of an inherited predisposition to colon cancer in a family with a variable number of adenomatous polyps. N Engl J Med 1990, **322**:904-908

Levi S, Urbono-Ispizua A, Gill R, Thomas DM, Gilbertson J, Foster C, Marshall CJ. Multiple K-ras codon 12 mutations in cholangiocarcinomas demonstrated with a sensitive polymerase chain reaction technique. Cancer Res 1991, **51**:3497-3502

Levine A. p53, the cellular gatekeeper for growth and division. Cell 1997, **88**:323-331

Levy DB, Smith KJ, Beazer-Barclay Y, Hamilton SR, Vogelstein B, Kinzler KW. Inactivation of both APC alleles in human and mouse tumors. Cancer Res 1994, **54**:5953-5958

Leyns L, Bouwmeester T, Kim SH, Piccolo S, De Robertis EM. Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann organizer. Cell 1997, **88**:747-756

Li FP, Fraumeni JF,Mulvihi IJJ, Blattner WA, Dreyfus MG, Tucker MA, Miller RW. A cancer family syndrome in 24 kindreds. Cancer Res 1988, **48**:5358-5362

Li ZH, Zheng J, Weiss LM, Shibata D. c-k-ras and p53 mutations occur very early in adenocarcinoma of the lung. Am J Pathol 1994a, **144**:303-309

Li Y, Nichols MA, Shay JW, Xiong Y. Transcriptional repression of the D-type cyclin-dependent kinase inhibitor p16 by the retinoblastoma susceptibility gene product pRb. Cancer Res 1994b, **54**:6078-6082

Lijam N, Paylor R, McDonald MP, Crawley JN, Deng CX, Herrup K, Stevens KE, Maccaferri G, McBain CJ, Sussman DJ, Wynshaw-Boris A. Social interaction and sensorimotor gating abnormalities in mice lacking Dvl1. Cell 1997, **90**:895-905

Lin R, Thompson S, Priess JR. pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos. Cell 1995, **83**:599-609

Lin L, Sahr KE, Chishti AH. Identification of the mouse homologue of human discs large and rat SAP97 genes. Biochim Biophys Acta 1997, **1362**:1-5

Liu C, Tsao MS. Proto-oncogene and growth factor/receptor expression in the establishment of primary human non-small cell lung carcinoma cell lines. Am J Pathol 1993, **142**:413-423

Lochter A, Srebrow A, Sympson CJ, Terracio N, Werb Z, Bissell MJ. Misregulation of stromelysin-1 expression in mouse mammary tumor cells accompanies acquisition of stromelysin-1-dependent invasive properties. J Biol Chem 1997, **272**:5007-5015

London SJ, Daly AK, Leathart JBS, Navidi WC, Carpenter CC, Idle JR. Genetic polymorphism of CYP2D6 and lung cancer risk in African-Americans and Cucasians in Los Angeles County. Carcinogenesis 1997, **18**:1203-1214

Lopez-Ilasaca M, Crespo P, Pellici PG, Gutkind JS, Wetzker R. Linkage of G protein-coupled receptors to the MAPK signaling pathway through PI 3-kinase gamma. Science 1997, **275**:394-397

Losi L, Roncucci L, di Gregorio C, de Leon MP, Benhattar J. K-ras and p53 mutations in human colorectal aberrant crypt foci. J Pathol 1996, **178**:259-263

Lowy DR, Willumsen BM. Function and regulation of ras. Annu Rev Biochem 1993, **62**:851-891

Lloyd JW Long-term mortality study of steelworkers. V. Respiratory cancer in coke plant workers. J Occup Med 1971, **13**:53-68

Lue RA, Marfatia SM, Branton D, Chishti AH. Cloning and characterization of hdlg: the human homologue of the Drosophila discs large tumor suppressor binds to protein 4.1. Proc Natl Acad Sci U S A 1994, **91**:9818-9822

van der Luijt RB, Meera Khan P, Vasen HF, Breukel C, Tops CM, Scott RJ, Fodde R Germline mutations in the 3' part of APC exon 15 do not result in truncated proteins and are associated with attenuated adenomatous polyposis coli. Hum Genet 1996, **98**:727-734

Lukas J, Parry D, Aagaard L, Mann DJ, Bartkova J, Strauss M, Peters G, Bartek J. Retinoblastoma protein dependent cell-cycle inhibition by the tumour suppressor p16. Nature 1995, **375**:503-506

Luongo C, Moser AR, Gledhill S, Dove WF. Loss of Apc+ in intestinal adenomas from Min mice. Cancer Res 1994, **54**:5947-5952

Luongo C, Dove WF. Somatic genetic events linked to the Apc locus in intestinal adenomas of the Min mouse. Genes Chromosomes Cancer 1996, **17**:194-198

Lutterbach B, Hann SR. Hierarchical phosphorylation at N-terminal transformation-sensitive sites in c-Myc protein is regulated by mitogens and in mitosis. Mol Cell Biol 1994,**14**:5510-5522

Lutterbach B, Hann SR. Overexpression of c-Myc and cell immortalization alters c-Myc phosphorylation. Oncogene 1997, **14**:967-975

Luttges J, Schlehe B, Menke MA, Vogel I, Henne-Bruns D, Kloppel G. The K-ras mutation pattern in pancreatic ductal adenocarcinoma usually is identical to that in associated normal, hyperplastic, and metaplastic ductal epithelium. Cancer 1999, **85**:1703-1710

Macara IG, Lounsbury KM, Richards SA, McKiernan C, Bar-Sagi D. The Ras superfamily of GTPases. FASEB J 1996, **10**:625-630

MacPhee M, Chepenik KP, Liddell RA, Nelson KK, Siracusa LD, Buchberg AM. The secretory phospholipase A2 gene is a candidate for the Mom1 locus, a major modifier of APC ^{Min} induced intestinal neoplasia. Cell 1995, **81**:957-966

Madri JA and Carter D. Scar cancers of the lung: origin and significance. Hum Pathol 1984, **15**:625-631

Mahmoud NN, Boolbol SK, Bilinski RT, Martucci C, Chadburn A, Bertagnolli MM. Apc gene mutation is associated with a dominant-negative effect upon cell migration. Cancer Res 1997, **57**:5045-5050

Makino K, Kuwahara H, Masuko N, Nishiyama Y, Morisaki T, Sasaki J, Nakao M, Kuwano A, Nakata M, Ushio Y, Saya H. Cloning and characterization of NE-dlg: a novel human homolog of the Drosophila discs large (dlg) tumor suppressor protein interacts with the APC protein. Oncogene 1997, **14**:2425-2433

Malkin D, Li FP, Strong LC, Fraumeni JF, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA, Freind SH. Germline p53 mutations in a familial syndrome of breast cancer, sarcomas and other neoplasms. Science 1990, **250**:1233-1238

Malkin D. P53 and the Li-Fraumeni syndrome. Cancer Genet Cytogenet 1993, **66**:83-92

Malkin D. Germline p53 mutations and heritable cancer. Annu Rev Genet 1994, **28**:443-465

Malumbres M, Pellicer A. RAS pathways to cell cycle control and cell transformation. Front Biosci 1998, **3**:d887-912

Mao L, Merlo A, Bedi G, Shapiro GI, Edwards CD, Rollins BJ, Sidransky D. A novel p16INK4A transcript. Cancer Res 1995, **55**:2995-2997

Marais R, Wynne J, Treisman R. The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. Cell 1993, **73**:381-393

Marchetti A, Doglioni C, Barbareschi M, Buttitta F, Pellegrini S, Gaeta P, La Rocca R, Merlo G, Chella A, Angeletti CA, Dalla Palma P, Bevilacqua G. Cyclin D1 and retinoblastoma susceptibility gene alterations in non-small cell lung cancer. Int J Cancer 1998, **75**:187-192

Marcus EA, Kintner C, Harris W. The Role of GSK3beta in Regulating Neuronal Differentiation in Xenopus laevis. Mol Cell Neurosci 1998, **12**:269-280

Marshall CJ. Ras effectors. Curr Opin Cell Biol 1996, **8**:197-204

Marcu KB, Bossone SA, Patel AJ. myc function and regulation. Annu Rev Biochem 1992, **61**:809-860

Marrs JA, Nelson WJ. Cadherin cell adhesion molecules in differentiation and embryogenesis. Int Rev Cytol 1996, **165**:159-205

Matrisian LM. The matrix-degrading metalloproteinases. Bioessays 1992, **14**:455-463

Matsubayashi H, Watanabe H, Yamaguchi T, Ajioka Y, Nishikura K, Iwafuchi M, Yamano M, Kijima H, Saito T. Multiple K-ras mutations in hyperplasia and carcinoma in cases of human pancreatic carcinoma. Jpn J Cancer Res 1999, **90**:841-848

Matsumine A, Ogai A, Senda T, Okumura N, Satoh K, Baeg GH, Kawahara T, Kobayashi S, Okada M,Toyoshima K, Akiyama T. Binding of APC to the human homolog of the Drosophila discs large tumor suppressor protein. Science 1996, **272**:1020-1023

Matsumoto K, Asano T, Endo T. Novel small GTPase M-Ras participates in reorganization of actin cytoskeleton. Oncogene 1997, **15**:2409-2417

McBurney MW, Sutherland LC, Adra CN, Leclair B, Rudnicki MA, Jardine K. The mouse Pgk-1 gene promoter contains an upstream activator sequence. Nucleic Acids Res 1991, **19**:5755-5761

McBurney MW, Staines WA, Boekelheide K, Parry D, Jardine K, Pickavance L. Murine PGK-1 promoter drives widespread but not uniform expression in transgenic mice. Dev Dyn 1994 , **200**:278-293

McCrea PD, Turck CW, Gumbiner B. A homolog of the armadillo protein in Drosophila (plakoglobin) associated with E-cadherin. Science 1991, **254**:1359-1361

McDonnell S, Navre M, Coffey RJ Jr, Matrisian LM. Expression and localization of the matrix metalloproteinase pump-1 (MMP-7) in human gastric and colon carcinomas. Mol Carcinog 1991, **4**:527-33

McKendry R, Hsu SC, Harland RM, Grosschedl R. LEF-1/TCF proteins mediate wnt-inducible transcription from the Xenopus nodal-related 3 promoter. Dev Biol 1997, **192**:420-431

McMahon AP, Gavin BJ, Parr B, Bradley A, McMahon JA. The Wnt family of cell signalling molecules in postimplantation development of the mouse. Ciba Found Symp 1992;**165**:199-212; discussion 212-218

MacPhee M, Chepenik KP, Liddell RA, Nelson KK, Siracusa LD, Buchberg AM. The secretory phospholipase A2 gene is a candidate for the Mom1 locus, a major modifier of ApcMin-induced intestinal neoplasia. Cell 1995, **81**:957-966

McWilliams JE, Sanderson BJ, Harris EL Richert-Boe KE, Henner WD. Glutathione S-transferase M1 (GSTM1) deficiency and lung cancer risk. Cancer Epidemiol Biomarkers Prev 1995, **4**:589-594

Merlo A, Herman JG, Mao L, Lee DJ, Gabrielson E, Burger PC, Baylin SB, Sidransky D. 5'CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS in human cancers. Nature Med 1995, **1**:686-692

Metzger D, Clifford J, Chiba H, Chambon P. Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. Proc Natl Acad Sci *U S A* 1995, **92**:6991-6995

Meyer EC, Liebow AA. A relationship of interstitial pneumonia honeycombing and atypical epithelial proliferation to cancer of the lung. Cancer 1965, **18**:322-325

Meyer BI, Gruss P. Mouse Cdx-1 expression during gastrulation. Development 1993, **117**:191-203

Midgley CA, White S, Howitt R, Save V, Dunlop MG, Hall PA, Lane DP, Wyllie AH, Bubb VJ. APC expression in normal human tissues. J Pathol 1997, **181**:426-433

Miller R, Nelems B, Evans K, Muller N, Ostrow D. Glandular neoplasia of the lung. A proposed analogy to colonic tumours. Cancer 1988, **61**:1009-1014

Miller R. Bronchioloalveolar cell adenomas. The Am J Surg Pathol 1990, **14**:904-912

Miller RR. Alveolar atypical hyperplasia in association with primary pulmonary adenocarcinoma: a clinicopathological study of 10 cases. Thorax 1993, **48**:679-680

Mills NE, Fishman CL, Rom WN, Dubin N, Jacobson DR. Increased prevalence of K-ras oncogene mutations in lung adenocarcinoma. Cancer Res 1995, **55**:1444-1447

Miniou P, Tiziano D, Frugier T, Roblot N, Le Meur M, Melki J. Gene targeting restricted to mouse striated muscle lineage. Nucleic Acids Res 1999, **27**:e27

Minna J. The molecular biology of lung cancer pathogenesis. Chest 1993, **103**:449s-456s

Mishina T, Dosaka-Akita H, Kinoshita I, Hommura F, Morikawa T, Katoh H, Kawakami Y. Cyclin D1 expression in non-small-cell lung cancers: its association with altered p53 expression, cell proliferation and clinical outcome. Br J Cancer 1999, **80**:1289-1295

Mitsudomi T, Viallet J, Mulshine JL, Linnoila RI, Minna JD, Gazdar AF. Mutations of ras genes distinguish a subset of non-small-cell lung cancer cell lines from small-cell lung cancer cell lines. Oncogene 1991, **8**:1353-1362

Miyashiro I, Senda T, Matsumine A, Baeg GH, Kuroda T, Shimano T, Miura S, Noda T, Kobayashi S, Monden M,et al Subcellular localization of the APC protein: immunoelectron microscopic study of the association of the APC protein with catenin. Oncogene 1995, **11**:89-96

Miyoshi Y, Nishisho I, Miki Y, Mori T, Kinzler KW, Vogelstein B, Nakamura Y. Insertion/deletion polymorphism and other restriction fragment length polymorphism's in the MCC gene. Jpn J Cancer Res 1992a, **83**:10-14

Miyoshi Y, Nagase H, Ando H, Horii A, Ichii S, Nakatsuru S, Aoki T, Miki Y, Mori T, Nakamura Y. Somatic mutations of the *APC* gene in colorectal tumours: mutation cluster region in the APC gene. Hum Mol Genet 1992b, **4**:229-233

Mizuno T, Kaibuchi K, Yamamoto T, Kawamura M, Sakoda T, Fujioka H, Matsuura Y, Takai Y. A stimulatory GDP/GTP exchange protein for smg p21 is active on the post-translationally processed form of c-Ki-ras p21 and rhoA p21. Proc Natl Acad Sci U S A 1991, **88**:6442-6446

Molenaar M, van de Wetering M, Oosterwegel M, Peterson-Maduro J, Godsave S, Korinek V, Roose J, Destree O, Clevers H. XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. Cell 1996, **86:**391-399

Moodie SA, Paris M, Villafranca E, Kirshmeier P, Willumsen BM, Wolfman A. Different structural requirements within the switch II region of the Ras protein for interactions with specific downstream targets. Oncogene 1995, **11**:447-454

Morais Cabral JH, Petosa C, Sutcliffe MJ, Raza S, Byron O, Poy F, Marfatia SM, Chishti AH, Liddington RC. Crystal structure of a PDZ domain. Nature 1996, **382**:649-652

Mori M, Tezuka F, Chiba R, Funae Y, Watanabe M, Nukiwa T, Takahashi T. Atypical adenomatous hyperplasia and adenocarcinoma of the human lung. Their heterology in form and analogy in immunohistochemical characteristics. Cancer 1996, **77**:665-674

Morin PJ, Sparks AB, Korinek V, Barker N, Clevers H, Vogelstein B, Kinzler KW. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. Science 1997, **275**:1787-1790

Morrison DK, Cutler Jr RE. The complexity of Raf-1 regulation. Curr Opin Cell Biol 1997, **2**:174-179

Morrison EE, Wardleworth BN, Askham JM, Markham AF, Meredith DM. EB1, a protein which interacts with the APC tumour suppressor, is associated with the microtubule cytoskeleton throughout the cell. Oncogene 1998,**17**:3471-3477

Moser AR, Pitot HC, Dove WF. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. Science 1990, **247**:32232-4

Moser AR, Mattes EM, Dove WF, Lindstrom MJ, Haag JD, Gould MN. ApcMin, a mutation in the murine Apc gene, predisposes to mammary carcinomas and focal alveolar hyperplasias. *Proc Natl Acad Sci U S A* 1993, **90**:8977-8981

Moser AR, Shoemaker AR, Connelly CS, Clipson L, Gould KA, Luongo C, Dove WF, Siggers PH, Gardner RL. Homozygosity for the Min allele of Apc results in disruption of mouse development prior to gastrulation. Dev Dyn 1995, **203**:422-433

Mortensen RM, Zubiaur M, Neer EJ, Seidman JG. Embryonic stem cells lacking a functional inhibitory G-protein subunit (alpha i2) produced by gene targeting of both alleles. Proc Natl Acad Sci *U S A* 1991, **88**:7036-7040

Motokura T, Arnold A. Cyclin D and oncogenesis. Curr Opin Genet Dev 1993, **3**:5-10

Mountain CF. A new international staging system for lung cancer. Chest 1986, **89**:S225-S233

Mountain C. Revisions in the international staging system for staging lung cancer. Chest 1997, **111**:1710-1717

Muhua L, Adames NR, Murphy MD, Shields CR, Cooper JA. A cytokinesis checkpoint requiring the yeast homologue of an APC-binding protein. Nature 1998, **393**:487-491

Mukoyama M, Nakajima M, Horiuchi M, Sasamura H, Pratt RE, Dzau VJ. Expression cloning of type 2 angiotensin II receptor reveals a unique class of seven-transmembrane receptors. J Biol Chem 1993, **268**:24539-24542

Muller H, Lukas J, Schneider A, Warthoe P, Bartek J, Eilers M, Strauss M. Cyclin D1 expression is regulated by the retinoblastoma protein. Proc Natl Acad Sci U S A 1994, **91**:2945-2949

Müller BM, Kistner U, Veh RW, Cases-Langhoff C, Becker B, Gundelfinger ED, Garner CC. Molecular characterization and spatial distribution of SAP97, a novel presynaptic protein homologous to SAP90 and the Drosophila discs-large tumor suppressor protein. J Neurosci 1995, **15:**2354-2366

Munemitsu S, Souza B, Muller O, Albert I, Rubinfeld B, Polakis P. The APC gene product associates with microtubules in vivo and promotes their assembly in vitro. Cancer Res 1994, **54**:3676-3681

Munemitsu S, Albert I, Souza B, Rubinfeld B, Polakis P. Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. Proc Natl Acad Sci U S A 1995, **92**:3046-3050

Murre C, McCaw PS, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell 1989, **56**:777-783

Musgrove EA, Lee CS, Buckley MF, Sutherland RL. Cyclin D1 induction in breast cancer cells shortens G1 and is sufficient for cells arrested in G1 to complete the cell cycle. Proc Natl Acad Sci U S A 1994, **91**:8022-8026

Nagase H, Nakamura Y. Mutations of the APC (adenomatous polyposis coli) gene. Hum Mutation 1993, **2**:425-434

Nakachi K, Imai K, Hayashi S, Kawajiri K. Polymorphisms of the CYP1A1 and glutathione S-transferase genes associated with susceptability to lung cancer in relation to cigarette dose in a Japanese population. Cancer Res 1993, **53**:2994-2999

Nakagawa H, Murata Y, Koyama K, Fujiyama A, Miyoshi Y, Monden M, Akiyama T, Nakamura Y. Identification of a brain-specific APC homologue, APCL, and its interaction with beta-catenin. Cancer Res 1998, **58**:5176-5181

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, et al. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 1987, **235**:1616-1622

Nakamura T, Hamada F, Ishidate T, Anai K, Kawahara K, Toyoshima K, Akiyama T. Axin, an inhibitor of the Wnt signalling pathway, interacts with beta-catenin, GSK-3beta and APC and reduces the beta-catenin level. Genes Cells 1998, **3**:395-403

Nakanishi K. Alveolar epithelial hyperplasia and adenocarcinoma of the lung. Arch Pathol Lab Med 1990, **114**:363-368

Nakatsuru S, Yanagisawa A, Furukawa Y, Ichii S, Kato Y, Nakamura Y, Horii A. Somatic mutations of the APC gene in precancerous lesion of the stomach. Hum Mol Genet 1993, **2**:1463-1465

Nakatsuru S, Yanagisawa A, Ichii S, Tahara E, Kato Y, Nakamura Y, Horii A. Somatic mutation of the APC gene in gastric cancer: frequent mutations in very well differentiated adenocarcinoma and signet-ring cell carcinoma. Hum Mol Genet 1992, **1**:559-563

Nakayama H, Noguchi M, Tsuchiya R, Kodama T, Shimosato Y. Clonal growth of atypical adenomatous hyperplasia of the lung: cytofluorometric analysis of nuclear DNA content. Modern Pathol 1990, **3**:314-320

Nash AD, Stout AP. Giant cell carcinoma of the lung: report of oveet cases. Cancer 1958, **11**:369-376

Nasiell M, Auer G, Kato H. Cytological studies in man and animals on development of bronchogenic carcinomas. In lung carcinomas. Edited by E. McDowell New York, Churchill Livingstone.

Nesnow S, Mass MJ, Ross JA, Galati AJ, Lambert GR, Gennings C, Carter WH Jr, Stoner GD. Lung tumorigenic interactions in strain A/J mice of five environmental polycyclic aromatic hydrocarbons. Environ Health Perspect 1998 , **106** Suppl 6:1337-1346

Nassar N, Horn G, Herrmann C, Block C, Janknecht R, Wittinghofer A. Ras/Rap effector specificity determined by charge reversal. Nat Struct Biol 1996, **3**:723-729

Näthke IS, Adams CL, Polakis P, Sellin JH, Nelson WJ. The adenomatous polyposis coli tumor suppressor protein localizes to plasma membrane sites involved in active cell migration. Cell Biol 1996, **134**:165-179

Nees M, Homann N, Discher H, Andl T, Enders C, Herold-Mende C, Schuhmann A, Bosch FX. Expression of mutated p53 occurs in tumor-distant epithelia of head and neck cancer patients: a possible molecular basis for the development of multiple tumors. Cancer Res 1993, **53**:4189-4196

Nesbit CE, Tersak JM, Prochownik EV. MYC oncogenes and human neoplastic disease. Oncogene 1999, **18**:3004-3016

Nesnow S, Ross JA, Mass MJ, Stoner GD. Mechanistic relationships between DNA adducts, oncogene mutations, and lung tumorigenesis in strain A mice. Exp Lung Res 1998, **24**:395-405

Neufeld KL, White RL. Nuclear and cytoplasmic localizations of the adenomatous polyposis coli protein. Proc Natl Acad Sci U S A 1997, **94**:3034-3039

Neumann C, Cohen S. Morphogens and pattern formation. Bioessays 1997, **19**:721-729

Nevins JR. Toward an understanding of the functional complexity of the E2F and retinoblastoma families. cell Growth Differ 1998, **9**:585-593

Newell KJ, Witty JP, Rodgers WH, Matrisian LM. Expression and localization of matrix-degrading metalloproteinases during colorectal tumorigenesis. Mol Carcinog 1994, **10**:199-206

Niethammer M, Kim E, Sheng M. Interaction between the C terminus of NMDA receptor subunits and
multiple members of the PSD-95 family of membrane-associated guanylate kinases. J Neurosci 1996, **16**:2157-2163

Nigro JM, Baker SJ, Preisinger AC, Jessup JM, Hostetter R, Cleary K, Bigner SH, Davidson N, Baylin S, Devilee P, et al. Mutations in the p53 gene occur in diverse human tumour types. Nature 1989, **342**:705-708

Nishisho I, Nakamura Y, Miyoshi Y, Miki Y, Ando H, Horri A, Koyoma K, Utsonomya J, Baba S, Hedge P, Markham A, Krush AJ, Peterson G, Hamilton SR, Nilbert MC, Levy DB, Bryon TM, Presinger AC, Smith JK, Su L, Kinzler KW, Volgelstein B. Mutations of chromosome 5q21 genes in FAP patients and colorectal cancer patients. Science 1991, **253**:665-669

Nobori T, Miura K, Wu DJ, Lois A, Takabayashi K, Carson DA. Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. Nature 1994, **368**:753-756

Noguchi M, Maezawa N, Nakanishi Y, Matsuno Y, Shimosato Y, Hirohashi S. Application of the p53 gene mutation pattern for differential diagnosis of primary versus metastatic lung carcinomas. Diagn Mol Pathol 1993, **2**:29-35

Nowell PC. The clonal evolution of tumor cell populations. Science 1976, **194**:23-28

Nuorva K, Soini Y, Kamel D, Autio-Harmainen H, Risteli L, Risteli J, Vahakangas K, Paakko P. Concurrent p53 expression in bronchial dysplasias and squamous cell lung carcinomas. Am J Pathol 1993, **142**:725-732

Nusse R, Varmus HE. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell 1982, **31**:99-109

Nusse R, van Ooyen A, Cox D, Fung YK, Varmus H. Mode of proviral activation of a putative mammary oncogene (int-1) on mouse chromosome 15. Nature 1984, **307**:131-136

Nusse R, Varmus HE. Wnt genes. Cell 1992, **69**:1073-1087

Nusse R, Samos CH, Brink M, Willert K, Cadigan KM, Wodarz A, Fish M, Rulifson E. Cell culture and whole animal approaches to understanding signaling by Wnt proteins in Drosophila. Cold Spring Harb Symp Quant Biol 1997, **62**:185-190

Nyberg F, Agrenius V, Svartengren K, Svensson C, Pershagen G. Dietary factors and risk of lung cancer in never-smokers. nt J Cancer 1998a, **78**:430-436

Nyberg F. More about: saturated fat intake and lung cancer risk among nonsmoking women in Missouri. J Natl Cancer Inst 1998b, **90**:1104-1105

Nyberg F, Hou SM, Hemminiki K, Lambert B, Pershagen G. Glutathione S-transferase Mlu1 and N-acetyltransferase 2 genetic polymorphisms and exposure to tabacco smoke in non-smoking and smoking lung cancer patients and population controls. Cancer Epidemiol Biomarkers Prev 1998c, 7:875-883

Oda H, Imai Y, Nakatsuru Y, Hata J, Ishikawa T. Somatic mutations of the APC gene in sporadic hepatoblastomas. Cancer Res 1996, **56**:3320-3323

Ohshima S, Shimizu Y, Takahama M. Detection of c-Ki-ras gene mutation in paraffin sections of adenocarcinoma and atypical bronchioloalveolar cell hyperplasia of human lung. Virchows Arch 1994, **424**:129-134

Okamoto A, Hussain SP, Hagiwara K, Spillare EA, Rusin MR, Demetrick DJ, Serrano M, Hannon GJ, Shiseki M, Zariwala M, et al. Mutations in the p16INK4/MTS1/CDKN2, p15INK4B/MTS2, and p18 genes in primary and metastatic lung cancer. *Cancer Res* 1995, **55**:1448-1451

Old RW, Primrose SB. Principles of gene manipulation. An introduction to genetic engineering.Fourth edition 1989.

Olivero M, Rizzo M, Madeddu R, Casadio C, Pennacchietti S, Nicotra MR, Prat M, Maggi G, Arena N, Natali PG, Comoglio PM, Di Renzo MF. Overexpression and activation of hepatocyte growth factor/scatter factor in human non-small-cell lung carcinomas. Br J Cancer 1996, **74**:1862-1868

Olschwang S, Tiret A, Laurent-Puig P, Muleris M, Parc R, Thomas G. Restriction of ocular fundus lesions to a specific subgroup of APC mutations in adenomatous polyposis coli patients. Cell 1993, **75**::959-968

Oosterwegel M, van de Wetering M, Timmerman J, Kruisbeek A, Destree O, Meijlink F, Clevers H. Differential expression of the HMG box factors TCF-1 and LEF-1 during murine embryogenesis. Development 1993, **118**:439-448

Orban PC, Chui D, Marth JD. Tissue- and site-specific DNA recombination in transgenic mice. Proc Natl Acad Sci U S A 1992, **89**:6861-6865

Orford K, Crockett C, Jensen JP, Weissman AM, Byers SW. Serine phosphorylation-regulated ubiquitination and degradation of beta-catenin. J Biol Chem 1997, **272**:24735-2478

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T. Detection of polymorphisms of human DNA by gel electrophoresis as single strand conformational polymorphisms. Proc Natl Acad Sci. USA 1989, **86**:2766-2770

Orsulic S, Peifer M. Cell-cell signalling: Wingless lands at last. Curr Biol 1996, **6**:1363-1367

Orsulic S, Huber O, Aberle H, Arnold S, Kemler R. E-cadherin binding prevents beta-catenin nuclear localization and beta-catenin/LEF-1-mediated transactivation. J Cell Sci 1999, **112** :1237-1245

Oshima M, Takahashi M, Oshima H, Tsutsumi M, Yazawa K, Sugimura T, Nishimura S, Wakabayashi K, Taketo MM. Effects of docosahexaenoic acid (DHA) on intestinal polyp development in Apc delta 716 knockout mice. Carcinogenesis 1995, **16**:2605-2607

Packham G, Cleveland JL. c-Myc and apoptosis. Biochim Biophys Acta 1995, **1242**:11-28

Paget S. The distribution of secondary growths in cancer of the breast. Lancet 1989, **1**:571-573

Pajouh MS, Nagle RB, Breathnach R, Finch JS, Brawer MK, Bowden GT. Expression of metalloproteinase genes in human prostate cancer. J Cancer Res Clin Oncol 1991, **117**:144-150

Palay SL, Chan-Palay V. A guide to the synaptic analysis of the neuropil. Cold Spring Harb Symp Quant Biol 1976, **40**:1-16

Parkin DM, Muir CS, Whelon SL, Gao Y, Feslay J, Powell J. Cancer incidence in five continents. Lyon France International agency for research on cancer 1992.

Parkin DM, Pisoni P, Lopez AD, Masuyer E. Alast one in 7 cases of cancer is caused by smoking. Global estimates for 1995. Int J Cancer 1994, **59**:494-504

Parkhurst SM. Groucho: making its Marx as a transcriptional co-repressor. Trends Genet 1998, **14**:130-132

Parr BA, McMahon AP. Dorsalizing signal Wnt-7a required for normal polarity of D-V and A-P axes of mouse limb. Nature 1995, **374**:350-353

Pawson T. Protein-tyrosine kinases. Getting down to specifics. Nature 1995, **373**:477-478

Pegg AE, Dolan ME, Moschel RC. Structure, function, and inhibition of O6-alkylguanine-DNA alkyltransferase. Prog Nucleic Acid Res Mol Biol 1995, **51**:167-223

Peifer M, Wieschaus E. The segment polarity gene armadillo encodes a functionally modular protein that is the Drosophila homolog of human plakoglobin. Cell 1990, **63**:1167-1176

Peifer M, Berg S, Reynolds AB. A repeating amino acid motif shared by proteins with diverse cellular roles. Cell 1994, **76**:789-791

Pells S, Divjak M, Romanowski P, Impey H, Hawkins NJ, Clarke AR, Hooper ML, Williamson DJ. Developmentally-regulated expression of murine K-ras isoforms. Oncogene 1997, **15**:1781-1786

Perry WL 3rd, Vasicek TJ, Lee JJ, Rossi JM, Zeng L, Zhang T, Tilghman SM, Costantini F. Phenotypic and molecular analysis of a transgenic insertional allele of the mouse Fused locus. Genetics 1995, **141**:321-332

Peto R. Smoking and death: the past 40 years and the next 40. BMJ 1994, **309**:937-939

Pierce SB, Kimelman D. Regulation of Spemann organizer formation by the intracellular kinase Xgsk-3. Development 1995, **121**:755-765

Pizzuti A, Amati F, Calabrese G, Mari A, Colosimo A, Silani V, Giardino L, Ratti A, Penso D, Calza L, Palka G, Scarlato G, Novelli G, Dallapiccola B. cDNA characterization and chromosomal mapping of two human homologues of the Drosophila dishevelled polarity gene. Hum Mol Genet 1996, **5**:953-958

Plyte SE, Hughes K, Nikolakaki E, Pulverer BJ, Woodgett JR. Glycogen synthase kinase-3: functions in oncogenesis and development. Biochim Biophys Acta 1992, **1114**:147-162

Polakis P. The adenomatous polyposis coli (APC) tumor suppressor. Biochim Biophys Acta 1997, **1332**:F127-47

Ponting CP, Bork P. Pleckstrin's repeat performance: a novel domain in G-protein signaling? Trends Biochem Sci 1996, **21**:245-246

Ponting CP, Phillips C, Davies KE, Blake DJ. PDZ domains: targeting signalling molecules to sub-membranous sites. Bioessays 1997,**19**:469-479

Powell SM, Papadopoulos N, Kinzler KW, Smolinski KN, Meltzer SJ. APC gene mutations in the mutation cluster region are rare in esophageal cancers. Gastroenterology 1994, **107**:1759-1763

Powell WC, Matrisian LM. Complex roles of matrix metalloproteinases in tumor progression. Curr Top Microbiol Immunol 1996, **213**:1-21

Praml C, Savelyeva L, Le Paslier D, Siracusa LD, Buchberg AM, Schwab M, Amler LC. Human homologue of a candidate for the Mom1 locus, the secretory type II phospholipase A2 (PLA2S-II), maps to 1p35-36.1/D1S199. Cancer Res 1995, **55**:5504-5506

Prat M, Narsimhan RP, Crepaldi T, Nicotra MR, Natali PG, Comoglio PM. The receptor encoded by the human c-MET oncogene is expressed in hepatocytes, epithelial cells and solid tumors. Int J Cancer 1991, **49**:323-328

Prendergast GC, Ziff EB. DNA-binding motif. Nature 1989, **341**:392

Prendergast GC, Ziff EB. Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. Science 1991, **251**:186-189

Pretlow TP, Brasitus TA, Fulton NC, Cheyer C, Kaplan EL. K-ras mutations in putative preneoplastic lesions in human colon. J Natl Cancer Inst 1993,**85**:2004-2007

Prieve MG, Waterman ML. Nuclear localization and formation of beta-catenin-lymphoid enhancer factor 1 complexes are not sufficient for activation of gene expression. Mol Cell Biol 1999, **19**:4503-4515

Prins J, DeVries EG, Mulder NH. The myc family of oncogenes and their presence and importance in small-cell lung carcinoma and other tumour types. Anticancer Res 1993, **13**:1373-1385

Pronk GJ, Bos JL. The role of p21ras in receptor tyrosine kinase signalling. Biochim Biophys Acta 1994, **1198**:131-147

Pyles RB, Santoro IM, Groden J, Parysek LM. Novel protein isoforms of the APC tumor suppressor in neural tissue. Oncogene 1998, **16**:77-82

Quelle DE, Zindy F, Ashmun RA, Sherr CJ. Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. Cell 1995, **83**:993-1000

Rachwal WJ, Bongiorno PF, Orringer MB, Whyte RI, Ethier SP, Beer DG. Expression and activation of erbB-2 and epidermal growth factor receptor in lung adenocarcinomas. Br J Cancer 1995, **72**:56-64

Radford EP, Renard KG. Lung cancer in Swedish iron miners exposed to low doses of radon daughters. N Engl J Med 1984, **310**:1485-1494

Raeburn C and Spencer H. A study of the origin and development of lung cancer. Thorax 1953, **8**:1-10

Raino P. Ultrastructure of neoplastic cells in different histologic subtypes of pulmonary adenocarcinoma. Pathol Res Pract Commun 1983, **176**:216-235

Ranscht B. Cadherins and catenins: interactions and functions in embryonic development. Curr Opin Cell Biol 1994, 6:740-746

Rao SK, Fraire AE. Alveolar cell hyperplasia in association with adenocarcinoma of lung. Mod Pathol 1995, **8**:165-169

Ray MK, Fagan SP, Moldovan S, DeMayo FJ, Brunicardi FC. Beta cell-specific ablation of target gene using Cre-loxP system in transgenic mice. J Surg Res 1999, **84**:199-203

Reddy EP, Reynolds RK, Santos E, Barbacid M. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature 1982, **300**:149-152

Reed JC. Bcl-2 and the regulation of programmed cell death. J Cell Biol 1994, **124**:1-6

Reed JC. Bcl-2 family proteins. Oncogene 1998, **17**:3225-3236

Reichsman F, Smith L, Cumberledge S. Glycosaminoglycans can modulate extracellular localization of the wingless protein and promote signal transduction. J Cell Biol 1996, **135**:819-827

Reissmann PT, Koga H, Takahashi R, Figlin RA, Holmes EC, Piantadosi S, Cordon-Cardo C, Salamon DJ. Inactivation of the retinoblastoma susceptibility gene gene in non-small cell lung cancer. Oncogene 1993, **8**:1913-1919

Reissmann PT, Koga H, Figlin RA, Holmes EC, Slamon DJ. Amplification and overexpression of the cyclin D1 and epidermal growth factor receptor genes in non-small-cell lung cancer. Lung Cancer Study Group. Cancer Res Clin Oncol 1999, **125**:61-70

Renner C, Pfitzenmeier JP, Gerlach K, Held G, Ohnesorge S, Sahin U, Bauer S, Pfreundschuh M. RP1, a new member of the adenomatous polyposis coli-binding EB1-like gene family, is differentially expressed in activated T cells. J Immunol 1997, **159**:1276-1283

Reynolds SH, Anna CK, Brown KC, Wiest JS, Beattie EJ, Pero RW, Iglehart JD, Anderson MW. Activated protooncogenes in human lung tumors from smokers. Proc Natl Acad Sci U S A 1991, **88**:1085-1089

Reynolds AB, Daniel J, McCrea PD, Wheelock MJ, Wu J, Zhang Z. Identification of a new catenin: the tyrosine kinase substrate p120cas associates with E-cadherin complexes. Mol Cell Biol 1994, **14**:8333-8342

Riehl R, Johnson K, Bradley R, Grunwald GB, Cornel E, Lilienbaum A, Holt CE. Cadherin function is required for axon outgrowth in retinal ganglion cells in vivo. Neuron 1996, **17**:837-848

Riese J, Yu X, Munnerlyn A, Eresh S, Hsu SC, Grosschedl R, Bienz M. LEF-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic. Cell 1997, **88**:777-787

Riggleman B, Wieschaus E, Schedl P. Molecular analysis of the armadillo locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a Drosophila segment polarity gene. Genes Dev 1989, **3**:96-113

Rimm DL, Koslov ER, Kebriaei P, Cianci CD, Morrow JS. Alpha 1(E)-catenin is an actin-binding and –bundling protein mediating the attachment of F-actin to the membrane adhesion complex. Proc Natl Acad Sci U S A 1995, **92**:8813-8817

Rocheleau CE, Downs WD, Lin R, Wittmann C, Bei Y, Cha YH, Ali M, Priess JR, Mello CC. Wnt signaling and an APC-related gene specify endoderm in early C. elegans embryos. Cell 1997, **90**:707-716

Rodenhuis S, van de Wetering ML, Mooi WJ, Evers SG, van Zandwijk N, Bos JL. Mutational activation of the K-ras oncogene. A possible pathogenic factor in adenocarcinoma of the lung. N Engl J Med 1987, **317**:929-935

Rodehuis S, Slebos RJC, Boot AJ, Evers SG, Mooi WJ, Wagenaar SS, van Bodegom PC, Bos JL. Incidence and possible clinical significance of K-ras oncogene activation in adenocarcinoma of the lung. Cancer Res 1988, **48**:5738-5741

Rodenhuis S, Slebos RJC. Clinical significance of ras oncogene activation in human lung cancer. Cancer Res 1992, **52**:1731-1733

Rohlmann A, Gotthardt M, Willnow TE, Hammer RE, Herz J. Sustained somatic gene inactivation by viral transfer of Cre recombinase. Nat Biotechnol 1996, **14**:1562-1565

Roose J, Molenaar M, Peterson J, Hurenkamp J, Brantjes H, Moerer P, van de Wetering M, Destree O, Clevers H. The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. Nature 1998, **395**:608-612

Rosell R, Li S, Skacel Z, Mate JL, Maestre J, Canela M, Tolosa E, Armengol P, Barnadas A, Ariza A. Prognostic impact of mutated K-ras gene in surgically resected non-small cell lung cancer patients. Oncogene 1993, **8**:2407-2412

Rosell R, Monzo M, Pifarre A, Ariza A, Sanchez JJ, Moreno I, Maurel J, Lopez MP, Abad A, de Anta JM. Molecular staging of non-small cell lung cancer according to K-ras genotypes. Clin Cancer Res 1996, **2**:1083-1086

Rosell R, Molina F, Moreno I, Martinez E, Pifarre A, Font A, Li S, Skacel Z, Gomez-Codina J, Camps C, et al. Mutated K-ras gene analysis in a randomized trial of preoperative chemotherapy plus surgery versus surgery in stage IIIA non-small cell lung cancer. Lung Cancer 1995, **12** Suppl 1:S59-70

Rothberg PG, Spandorfer JM, Erisman MD, Staroscik RN, Sears HF, Petersen RO, Astrin SM. Evidence that c-myc expression defines two genetically distinct forms of colorectal adenocarcinoma. Br J Cancer 1985, **52**:629-632

Rubinfeld B, Souza B, Albert I, Muller O, Chamberlain SH, Masiarz FR, Munemitsu S, Polakis P. Association of the APC gene product with beta-catenin. Science 1993, **262**:1731-1734

Rubinfeld B, Souza B, Albert I, Munemitsu S, Polakis P. The APC protein and E-cadherin form similar but independent complexes with alpha-catenin, beta-catenin, and plakoglobin. J Biol Chem 1995, **270**:5549-5555

Rubinfeld B, Albert I, Porfiri E, Fiol C, Munemitsu S, Polakis P. Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. Science 1996, **272**:1023-1026

Rusch V, Baselga J, Cordon-Cardo C, Orazem J, Zaman M, Hoda S, McIntosh J, Kurie J, Dmitrovsky E. Differential expression of the epidermal growth factor receptor and its ligands in primary non-small cell lung cancers and adjacent benign lung. Cancer Res 1993, **53**(10 Suppl):2379-2385

Rusch V, Klimstra D, Linkov I, Dmitrovsky E. Aberrant expression of p53 or the epidermal growth factor receptor is frequent in early bronchial neoplasia and coexpression precedes squamous cell carcinoma development. Cancer Res 1995, **55**:1365-1372

Rusch V, Klimstra D, Venkatraman E, Pisters PW, Langenfeld J, Dmitrovsky E. Overexpression of the epidermal growth factor receptor and its ligand transforming growth factor alpha is frequent in resectable non-small cell lung cancer but does not predict tumor progression. Clin Cancer Res 1997, **3**:515-522

Ryan KM, Birnie GD. Cell-cycle progression is not essential for c-Myc to block differentiation. Oncogene 1997, **14**:2835-2843

Saam JR, Gordon JI. Inducible gene knockouts in the small intestinal and colonic epithelium. J Biol Chem 1999, **274**:38071-3882

Saarikoski ST, Voho A, Reinikainen M, Anttila S, Karjalainen A, Malaveille C, Vainio H, Husgafvel-Pursiainen K, Hirvonen A. Combined effect of polymorphic GST genes on individual susceptibility to lung cancer. Int J Cancer 1998, 77:516-521

Saccomanno G, Archer VE, Auerbach O, Saunders RP, Brennan LM. Development of carcinoma of the lung as reflected in exfoliated cells. Cancer 1974, **33**:256-270

Sakaguchi M, Fujii Y, Hirabayashi H, Yoon HE, Komoto Y, Oue T, Kusafuka T, Okada A, Matsuda H. Inversely correlated expression of p16 and Rb protein in non-small cell lung cancers. An immunohistochemical study. Int J Cancer 1996, **65**:442-445

Sakai K, Mitani K, Miyazaki J. Efficient regulation of gene expression by adenovirus vector-mediated delivery of the CRE recombinase. Biochem Biophys Res Commun 1995,**217**:393-401

Sakamuro D, Prendergast GC. New Myc-interacting proteins: a second Myc network emerges. Oncogene 1999, **18**:2942-2954

Sakanaka C, Weiss JB, Williams LT. Bridging of beta-catenin and glycogen synthase kinase-3beta by axin and inhibition of beta-catenin-mediated transcription. Proc Natl Acad Sci U S A 1998, **95**:3020-3023

Sakanaka C, Williams LT. Functional domains of axin. Importance of the C terminus as an oligomerization domain. J Biol Chem 1999, **274**:14090-14093

Sancar A. DNA excision repair. Annu Rev Biochem 1996, **65**:43-81

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 1977, **74**:5463-5467

Sanson B, White P, Vincent JP. Uncoupling cadherin-based adhesion from wingless signalling in Drosophila. Nature 1996, **383**:627-630

Santoro IM, Groden J. Alternative splicing of the APC gene and its association with terminal differentiation. Cancer Res 1997, **57**:488-494

Sato S, Nakamura Y, Tsuchiya E. Difference of allelotype between squamous cell carcinoma and adenocarcinoma of the lung. Cancer Res 1994, **54:**5652-5655

Sawada S, Littman DR. Identification and characterization of a T-cell-specific enhancer adjacent to the murine CD4 gene. Mol Cell Biol 1991, **11**:5506-5515

Sanz-Ortega J, Bryant B, Sanz-Esponera J, Asenjo JA, Saez MC, Torres A, Balibrea JL, Sobel ME, Merino MJ. LOH at the APC/MCC gene (5Q21) is frequent in early stages of non-small cell lung cancer. Pathol Res Pract 1999, **195**:677-680

Sauer B. Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae. Mol Cell Biol 1987, 7:2087-2096

Schaeffer BK, Glasner S, Kuhlmann E, Myles JL, Longnecker DS. Mutated c-K-ras in small pancreatic adenocarcinomas. Pancreas 1994 **9**:161-165

Schauer IE, Siriwardana S, Langan TA, Sclafani RA. Cyclin D1 overexpression vs. retinoblastoma inactivation: implications for growth control evasion in non-small cell and small cell lung cancer. Proc Natl Acad Sci U S A 1994, **91**:7827-7831

Scheele JS, Rhee JM, Boss GR. Determination of absolute amounts of GDP and GTP bound to Ras in mammalian cells: comparison of parental and Ras-overproducing NIH 3T3 fibroblasts. Proc Natl Acad Sci U S A 1995, **92**:1097-1100

Schenk PW, Snaar-Jagalska BE. Signal perception and transduction: the role of protein kinases. Biochim Biophys Acta 1999, **1449**:1-24

Schimke RT. Gene amplification in cultured cells. J Biol Chem 1988, **263**:5989-5992

Schlessinger J. The epidermal growth factor receptor rector as a multifunctional allosteric protein. Biochemistry, 1988, **27**:3119-3123

Schlessinger J. How receptor tyrosine kinases activate Ras.Trends Biochem Sci 1993, **18**:273-275

Schlessinger J, Lax I, Lemmon M. Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? Cell 1995, **83**:357-360

Schryver B, Hinck L, Papkoff J. Properties of Wnt-1 protein that enable cell surface association. Oncogene 1996, **13**:333-342

Schuuring E, Verhoeven E, Mooi WJ, Michalides RJ. Identification and cloning of two overexpressed genes, U21B31/PRAD1 and EMS1, within the amplified chromosome 11q13 region in human carcinomas. Oncogene 1992, **7**:355-361

Schwartz K, Richards K, Botstein D. BIM1 Encodes a Microtubule-binding Protein in Yeast. Mol Biol Cell 1997, **8**:2677-2691

Schwenk F, Baron U, Rajewsky K. A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. Nucleic Acids Res 1995, **23**:5080-5081

Schwenk F, Kuhn R, Angrand PO, Rajewsky K, Stewart AF. Temporally and spatially regulated somatic mutagenesis in mice. Nucleic Acids Res 1998, **26**:1427-1432

Scott RJ, van der Luijt R, Spycher M, Mary JL, Muller A, Hoppeler T, Haner M, Muller H, Martinoli S, Brazzola PL, et al. Novel germline APC gene mutation in a large familial adenomatous polyposis kindred displaying variable phenotypes. Gut 1995, **36**:731-736

Scott RJ, van der Luijt R, Spycher M, Mary JL, Muller A, Hoppeler T, Haner M, Muller H, Martinoli S, Brazzola PL, et al. Novel germline APC gene mutation in a large familial adenomatous polyposis kindred displaying variable phenotypes. Gut 1996, **38**:794

Scottish Health Statistics. National health service in Scotland, 1998.

Segev N, Cohen G. Control of circularization of bacteriophage P1 DNA in Escherichia coli. Virology 1981, **114**:333-342

Selbert S, Bentley DJ, Melton DW, Rannie D, Lourenco P, Watson CJ, Clarke AR. Efficient BLG-Cre mediated gene deletion in the mammary gland. Transgenic Res 1998, 7:387-396

Selikoff IJ. Lung cancer and mesothelioma during prospective surveillance of 1249 asbestos insulation workers, 1963-1974. Ann N Y Acad Sci 1976;**271**:448-456

Selikoff IJ, Hammond EC, Seidman H. Latency of asbestos disease among insulation workers in the United States and Canada. Cancer 1980, **46**:2736-2740

Senda T, Miyashiro I, Matsumine A, Baeg GH, Monden T, Kobayashil S, Monden M, Toyoshima K, Akiyama T. The tumor suppressor protein APC colocalizes with beta-catenin in the colon epithelial cells. Biochem Biophys Res Commun 1996, **223**:329-334

Senda T, Iino S, Matsushita K, Matsumine A, Kobayashi S, Akiyama T. Localization of the adenomatous polyposis coli tumour suppressor protein in the mouse central nervous system. Neuroscience 1998, **83**:857-866

Serrano M, Hannon GJ, Beach D. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK 4. Nature 1993, **336**:704-707

Shapiro GI, Edwards CD, Kobzik L, Godleski J, Richards W, Sugarbaker DJ, Rollins BJ. Reciprical Rb inactivation and p16INK4 expression in primary lung cancers and cell lines. Cancer Res 1995, **55**:505-509

Shaw GL, Falk RT, Deslauriers J, Frame JN, Nesbitt JC, Pass HI, Issaq HJ, Hoover RN, Tucker MA. Debrisoquine metabolism and lung cancer risk. Cancer Epidemiol Biomarkers Prev 1995, **4**:41-48

Shaw M, Cohen P, Alessi DR. Further evidence that the inhibition of glycogen synthase kinase-3beta by IGF-1 is mediated by PDK1/PKB-induced phosphorylation of Ser-9 and not by dephosphorylation of Tyr-216. FEBS Lett 1997, **416**:307-311

Sheffield VC, Beck JS, Kwitek AE, Sandstrom DW, Stone EM. The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. Genomics 1993, **16**:325-332

Sherr C. Cancer cell cycles. Science 1996, **274**:1672-1677

Shi D, He G, Cao S, Pan W, Zhang HZ, Yu D, Hung MC. Overexpression of the c-erbB-2/neu-encoded p185 protein in primary lung cancer. Mol Carcinog 1992, **5**:213-218

Shibamoto S, Hayakawa M, Takeuchi K, Hori T, Miyazawa K, Kitamura N, Johnson KR, Wheelock MJ, Matsuyoshi N, Takeichi M, et al. Association of p120, a tyrosine kinase substrate, with E-cadherin/catenin complexes. J Cell Biol 1995, **128**:949-957

Shibata H, Toyama K, Shioya H, Ito M, Hirota M, Hasegawa S, Matsumoto H, Takano H, Akiyama T, Toyoshima K, Kanamaru R, Kanegae Y, Saito I, Nakamura Y, Shiba K, Noda T. Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene. Science 1997, **278**:120-123

Shields PG, Bowman ED, Harrington AM, Doan VT, Weston A. Polycyclic aromatic hydrocarbon-DNA adducts in human lung and cancer susceptibility genes. Cancer Res 1993, **53**:3486-3492

Shirozu M, Tada H, Tashiro K, Nakamura T, Lopez ND, Nazarea M, Hamada T, Sato T, Nakano T, Honjo T. Characterization of novel secreted and membrane proteins isolated by the signal sequence trap method. Genomics 1996, **37**:273-280

Shimosato Y, Kodama T, Kameja T. Morphogenesis of peripheral type adenocarcinoma of the lung In: Shimosato Y, Melomed MR, Nettesheim P. editors. Morphogenesis of lung cancer, vol1 Boca Raton: CRC press 1982: 65-89

Shivapurkar N, Huang L, Ruggeri B, Swalsky PA, Bakker A, Finkelstein S, Frost A, Silverberg S. K-ras and p53 mutations in aberrant crypt foci and colonic tumors from colon cancer patients. Cancer Lett 1997, **115**:39-46

Shtutman M, Zhurinsky J, Simcha I, Albanese C, D'Amico M, Pestell R, Ben-Ze'ev A. The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. Proc Natl Acad Sci U S A 1999, **96**:5522-5527

Sikora K, Chan S, Evan G, Gabra H, Markham N, Stewart J, Watson J. c-myc oncogene expression in colorectal cancer. Cancer 1987, **59**:1289-1295

Simonds WF. G protein regulation of adenylate cyclase. Trends Pharmacol Sci 1999, **20**:66-73

Singer B, Hang B. What structural features determine repair enzyme specificity and mechanism in chemically modified DNA? Chem Res Toxicol 1997, **10:**713-732

Singh G, Katyal SL, Ordonez NG, Dail DH, Negishi Y, Weech VW, Marcus PB, Weldon-Linne CM, Axiotis CA, Alvarez-Fernandeze *et al.* Type II pneumocytes in pulmonary tumours. Implications for histogenesis. Arch. Path. Lab. Med. 1984, **108**:44-48

Singh-Kaw P, Zarnegar R, Siegfried JM. Stimulatory effects of hepatocyte growth factor on normal and neoplastic human bronchial epithelial cells. Am J Physiol 1995, **268**(6 Pt 1):L1012-20

Slaughter DP, Southwick HW,, Smejkal W. Field cancerization in oral stratified squamous epithelim:Clinical implications of multicentric origin. Cancer 1953, **6**:963-968

Slebos RJ, Evers SG, Wagenaar SS, Rodenhuis S. Cellular protooncogenes are infrequently amplified in untreated non-small cell lung cancer. Br J Cancer 1989, **59**:76-80

Slebos RJ, Kibbelaar RE, Dalesio O, Kooistra A, Stam J, Meijer CJ, Wagenaar SS, Vanderschueren RG, van Zandwijk N, Mooi WJ *et al.* K-ras oncogene activation as a prognostic marker in adenocarcinoma of the lung. N Engl J Med 1990, **323**:561-565

Slebos RJ, Habets GG, Evers SG, Mooi WJ, Rodenhuis S. Allele-specific detection of K-ras oncogene expression in human non-small-cell lung carcinomas. Int J Cancer 1991a, **48**:51-56

Slebos RJ, Hruban RH, Dalesio O, Mooi WJ, Offerhaus GJ, Rodenhuis S. Relationship between K-ras oncogene activation and smoking in adenocarcinoma of the human lung. J Natl Cancer Inst 1991b, **83**:1024-1027

Slebos R, Rodenhuis S. The ras gene family in human non-small cell lung cancer. J Natl Cancer Inst Manogr 1992, **13**:23-29

Slebos RJ, Baas IO, Clement MJ, Offerhaus GJ, Askin FB, Hruban RH, Westra WH. p53 alterations in atypical alveolar hyperplasia of the human lung. Hum Pathol 1998, **29**:801-808

Slusarski DC, Corces VG, Moon RT. Interaction of Wnt and a Frizzled homologue triggers G-protein-linked phosphatidylinositol signalling. Nature 1997, **390**:410-413

Smalley MJ, Sara E, Paterson H, Naylor S, Cook D, Jayatilake H, Fryer LG, Hutchinson L, Fry MJ, Dale TC. Interaction of axin and Dvl-2 proteins regulates Dvl-2-stimulated TCF-dependent transcription. EMBO J 1999, **18**:2823-2835

Smith AE. Viral vectors in gene therapy. Annu Rev Microbiol 1995, **49**:807-838

Smith DI, Huang H, Wang L. Common fragile sites and cancer. Int J Oncol 1998, **12**:187-196

Smith KJ, Johnson KA, Bryan TM, Hill DE, Markowitz S, Willson JK, Paraskeva C, Petersen GM, Hamilton SR,Vogelstein B, Kinzler KW. The APC gene product in normal and tumor cells. Proc Natl Acad Sci U S A 1993, **90**:2846-2850

Smith KJ, Levy DB, Maupin P, Pollard TD, Vogelstein B, Kinzler KW. Wild-type but not mutant APC associates with the microtubule cytoskeleton. Cancer Res 1994, **54**:3672-3675

Smith WC, McKendry R, Ribisi S Jr, Harland RM. A nodal-related gene defines a physical and functional domain within the Spemann organizer. Cell 1995, **82**:37-46

Smits R, van der Houven van Oordt W, Luz A, Zurcher C, Jagmohan-Changur S, Breukel C, Khan PM, Fodde R. Apc1638N: a mouse model for familial adenomatous polyposis-associated desmoid tumors and cutaneous cysts. Gastroenterology 1998, **114**:275-283

Smits R, Kielman MF, Breukel C, Zurcher C, Neufeld K, Jagmohan-Changur S, Hofland N, van Dijk J, White R, Edelmann W, Kucherlapati R, Khan PM, Fodde R. Apc1638T: a mouse model delineating critical domains of the adenomatous polyposis coli protein involved in tumorigenesis and development. Genes Dev 1999, **13**:1309-1321

Songyang Z, Fanning AS, Fu C, Xu J, Marfatia SM, Chishti AH, Crompton A, Chan AC, Anderson JM, Cantley LC. Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. Science 1997, **275**:73-77

Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 1975, **98:**503-517

Sozzi G, Miozzo M, Donghi R, Pilotti S, Coriani CT, Pastorino U, Della Porta G, Pierotti MA. Deletions of 17 p and p53 mutations in preneoplastic lesions of the lung. Cancer Res 1992, **52**:6079-6082

Sozzi G, Miozzo M, Pastorino U, Pilotti S, Donghi R, Giarola M, De Gregorio L, Manenti G, Radice P, Minoletti F, et al. Genetic evidence for an independent origin of multiple preneoplastic and neoplastic lung lesions. Cancer Res 1995, **55**:135-140

Sparks AB, Morin PJ, Vogelstein B, Kinzler KW Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. Cancer Res 1998, **58**:1130-1134

Staddon JM, Smales C, Schulze C, Esch FS, Rubin LL. p120, a p120-related protein (p100), and the cadherin/catenin complex. J Cell Biol 1995, **130**:369-381

Stark K, Vainio S, Vassileva G, McMahon AP. Epithelial transformation of metanephric mesenchyme in the developing kidney regulated by Wnt-4. Nature 1994, **372**:679-683

Stec DE, Davisson RL, Haskell RE, Davidson BL, Sigmund CD. Efficient liver-specific deletion of a floxed human angiotensinogen transgene by adenoviral delivery of Cre recombinase in vivo. J Biol Chem 1999, **274**:21285-2190

Steele RJC, Thompson AM, Hall PA, Lane DP. The p53 tumour suppressor gene. B J Surg 1998, **85**:1460-1467

Steitz SA, Tsang M, Sussman DJ. Wnt-mediated relocalization of dishevelled proteins. In Vitro Cell Dev Biol Anim 1996, **32**:441-445

Sternberg N. Bacteriophage P1 site-specific recombination. III. Strand exchange during recombination at lox sites. J Mol Biol 1981, **150**:603-608

Stewart M, Murphy C, Fristrom JW. The recovery and preliminary characterization of X chromosome mutants affecting imaginal discs of Drosophila melanogaster. Dev Biol 1972, **27**:71-83

Stirzaker C, Millar DS, Paul CL, Warnecke PM, Harrison J, Vincent PC, Frommer M, Clark SJ. Extensive DNA methylation spanning the Rb promoter in retinoblastoma tumors. Cancer Res 1997, **57**:2229-2237

Stokoe D, Macdonald SG, Cadwallader K, Symons M, Hancock JF. Activation of Raf as a result of recruitment to the plasma membrane. Science 1994, **264**:1463-1467

Stone S, Dayananth P, Jiang P, Weaver-Feldhaus JM, Tavtigian SV, Cannon-Albright L, Kamb A. Genomic structure, expression and mutational analysis of the P15 (MTS2) gene. Oncogene 1995b,**11**:987-991

Stone S, Jiang P, Dayananth P, Tavtigian SV, Katcher H, Parry D, Peters G, Kamb A. Complex structure and regulation of the p16(MTS1) locus. Cancer Res 1995a, **55**:2988-2994

Stott FJ, Bates S, James MC, McConnell BB, Starborg M, Brookes S, Palmero I, Ryan K, Hara E, Vousden KH, Peters G. The alternative product from the human CDKN2A locus, p14(ARF), participates in a regulatory feedback loop with p53 and MDM2. EMBO J 1998, **17**:5001-5014

Strauss M, Lukas J, Bartek J. Unrestricted cell cycling and cancer. Nat Med 1995, **1**:1245-1246

Su LK, Kinzler KW, Vogelstein B, Preisinger AC, Moser AR, Luongo C, Gould KA, Dove WF. Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. Science 1992, **256** : 668-70

Su LK, Vogelstein B, Kinzler KW. Association of the APC tumor suppressor protein with catenins. Science 1993, **262**:1734-1737

Su LK, Burrell M, Hill DE, Gyuris J, Brent R, Wiltshire R, Trent J, Vogelstein B, Kinzler KW. APC binds to the novel protein EB1. Cancer Res 1995, **55**:2972-2977

Sugimura H, Suzuki I, Hamada GS, Iwase T, Takahashi T, Nagura K, Iwata H, Watanabe S, Kino I, Tsugane S. Cytochrome P-450 lA1 genotype in lung cancer patients and controls in Rio de Janeiro, Brazil. Cancer Epidemiol Biomarkers Prev 1994, **3**:145-148

Sugio K, Ishida T, Yokoyama H, Inoue T, Sugimachi K, Sasazuki T. ras gene mutations as a prognostic marker in adenocarcinoma of the human lung without lymph node metastasis. Cancer Res 1992, **52**:2903-2906

Sugio K, Kishimoto Y, Virmani AK, Hung JY, Gazdar AF. K-ras mutations are a relatively late event in the pathogenesis of lung carcinomas. Cancer Res 1994, **54**:5811-5815

Sugio K, Molberg K, Albores-Saavedra J, Virmani AK, Kishimoto Y, Gazdar AF. K-ras mutations and allelic loss at 5q and 18q in the development of human pancreatic cancers. Int J Pancreatol 1997, **21**:205-217

Sundaresan V, Ganly P, Hasleton P, Rudd R, Sinha G, Bleehen NM, Rabbitts P. p53 and chromosome 3 abnormalities, characteristic of malignant lung tumours, are detectable in preinvasive lesions of the bronchus. Oncogene 1992, 7:1989-1997

Sunderman FW. Mechanisms of metal carcinogenesis. Environ. Health Perspect 1979, **40**:131-141

Sussman DJ, Klingensmith J, Salinas P, Adams PS, Nusse R, Perrimon N. Isolation and characterization of a mouse homolog of the Drosophila segment polarity gene dishevelled. Dev Biol 1994, **166**:73-86

Sutherland LC, St-Arnaud R, McBurney MW. An upstream activator sequence regulates the murine Pgk-1 promoter and binds multiple nuclear proteins. Gene Expr 1995, **4**:265-279

Suzuki Y, Orita M, Shiraishi M, Hayashi K, Sekiya T. Detection of RAS gene mutations in human lung cancers by single strand conformation polymorphism analysis of polymerase chain reaction products. Oncogene 1990, **5**:1037-1043

Tabata T, Fujimori T, Maeda S, Yamamoto M, Saitoh Y. The role of Ras mutation in pancreatic cancer, precancerous lesions, and chronic pancreatitis. Int J Pancreatol 1993, **14**:237-244

Tabin CJ, Bradley SM, Bargmann CI, Weinberg RA, Papageorge AG, Scolnick EM, Dhar R, Lowy DR, Chang EH. Mechanism of activation of a human oncogene. Nature 1982, **300**:143-149

Tagawa M, Murata M, Kimura H. Prognostic value of mutations and a germ line polymorphism of the p53 gene in non-small cell lung carcinoma: association with clinicopathological features. Cancer Lett 1998, **128**:93-99

Taguchi T, Cheng GZ, Bell DW, Balsara B, Liu Z, Siegfried JM, Testa JR. Combined chromosome microdissection and comparative genomic hybridization detect multiple sites of amlification DNA in a human lung carcinoma cell line. Genes Chrom Cancer 1997, **20**:208-212

Takada S, Stark KL, Shea MJ, Vassileva G, McMahon JA, McMahon AP. Wnt-3a regulates somite and tailbud formation in the mouse embryo. Genes Dev 1994, **8**:174-189

Takahashi T, Nau MM, Chiba I, Birrer MJ, Rosenberg RK, Vinocour M, Levitt M, Pass H, Gazdar AF, Minna JD. p53:a frequent target for genetic abnormalities in lung cancer. Science 1989, **246**:491-494

Takeichi M. Morphogenetic roles of classic cadherins. Curr Opin Cell Biol 1995, **7**:619-627

Tam SW, Stay JW, Pagano M. Differential expression and cell cycle regulation of the cyclin-dependent kinase 4 inhibitor p16 INK4. Cancer Res 1994, **54**:5816-5820

Tamura G, Maesawa C, Suzuki Y, Tamada H, Satoh M, Ogasawara S, Kashiwaba M, Satodate R. Mutations of the APC gene occur during early stages of gastric adenoma development. Cancer Res 1994, 54:1149-1151

Tanka H, Fujii Y, Hirabayashi H, Miyoshi S, Sakaguchi M, Yoon HE, Matsuda H. Disruption of the RB pathway and cell-proliferative activy in non-small cell lung cancers. Int J Cancer 1998, **79**:111-115

Tarutani M, Itami S, Okabe M, Ikawa M, Tezuka T, Yoshikawa K, Kinoshita T, Takeda J. Tissue-specific knockout of the mouse Pig-a gene reveals important roles for GPI-anchored proteins in skin development. Proc Natl Acad Sci *U S A* 1997, **94**:7400-7405

Tateishi M, Ishida T, Mitsudomi T, Kaneko S, Sugimachi K. Prognostic value of c-erbB-2 protein expression in human lung adenocarcinoma and squamous cell carcinoma. Eur J Cancer 1991, **27**:1372-1375

Testa J. Chromosome alterations in human lung cancer, in Mitchell J, Johnson D, Turrisi A (eds): Lung cancer:Principles and practice. Philadelphia PA, Lippincott-Raven 1996, pp55-71

Tetsu O, McCormick F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature 1999, **398**:422-426

The Scottish Office, Department of Health: Health in Scotland 1997

Thliveris A, Samowitz W, Matsunami N, Groden J, White R. Demonstration of promoter activity and alternative splicing in the region 5' to exon 1 of the APC gene. Cancer Res 1994, **54**:2991-2995

Thomas KR, Capecchi MR. Targeted disruption of the murine int-1 proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. Nature 1990, **346**:847-850

Thomas M, Kalita A, Azouly E, Labrecque S, Pim D, Banks L Matlashewski . Characterisation of two polymorphic forms of wildtype p53. 9th p53 workshop Crete Greece 1998.

Thompson AM, Morris RG, Wallace M, Wyllie AH, Steel CM, Carter DC. Allele loss from 5q21 (APC/MCC) and 18q21 (DCC) and DCC mRNA expression in breast cancer. Br J Cancer 1993, **68**:64-68

Tomlinson IP, Neale K, Talbot IC, Spigelman AD, Williams CB, Phillips RK, Bodmer WF. A modifying locus for familial adenomatous polyposis may be present on chromosome 1p35-p36. J Med Genet 1996a, **33** :268-273

Tomlinson IP, Beck NE, Neale K, Bodmer WF. Variants at the secretory phospholipase A2 (PLA2G2A) locus: analysis of associations with familial adenomatous polyposis and sporadic colorectal tumours. Ann Hum Genet 1996b, **60** :369-376

Torres MA, Yang-Snyder JA, Purcell SM, DeMarais AA, McGrew LL, Moon RT. Activities of the Wnt-1 class of secreted signaling factors are antagonized by the Wnt-5A class and by a dominant negative cadherin in early Xenopus development. J Cell Biol 1996, **133**:1123-1137

Torres RM, Kuhn R Laboratory protocols for conditional gene targeting 1997. Oxford University Press.

Townes TM, Lingrel JB, Chen HY, Brinster RL, Palmiter RD. Erythroid-specific expression of human beta-globin genes in transgenic mice. EMBO J 1985, **4**:1715-1723

Trahair JF, Neutra MR, Gordon JI. Use of transgenic mice to study the routing of secretory proteins in intestinal epithelial cells: analysis of human growth hormone compartmentalization as a function of cell type and differentiation. J Cell Biol 1989, **109**:3231-3242

Traub P. Intermediate filaments (Springer, Berlin.) 1985

Travis A, Amsterdam A, Belanger C, Grosschedl R. LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function. Genes Dev 1991, **5**:880-894

Tsang M, Lijam N, Yang Y, Beier DR, Wynshaw-Boris A, Sussman DJ. Isolation and characterization of mouse dishevelled-3. Dev Dyn 1996, **207**:253-262

Tsao MS, Liu N, Chen JR, Pappas J, Ho J, To C, Viallet J, Park M, Zhu H. Differential expression of Met/hepatocyte growth factor receptor in subtypes of non-small cell lung cancers. Lung Cancer 1998, **20**:1-16

Tsien JZ, Chen DF, Gerber D, Tom C, Mercer EH, Anderson DJ, Mayford M, Kandel ER, Tonegawa S. Subregion- and cell type-restricted gene knockout in mouse brain. Cell 1996, **87**:1317-1326

Tsuchiya E, Nakamura Y, Weng SY, Nagasowa K, Tsuchiya S, Sugano H, Kitagawa T. Allelotype of non-small cell lung carcinoma: comparison between loss of heterozygosity in squamous cell carcinoma and adenocarcinoma. Cancer Res 1992, **52**:2478-2481

Tsujimoto Y, Finger L, Yunis J et al Cloning of the chromosome breakpoint of neoplastic B cells with the t(14:18) chromosome translocation. Science 1984, **226**:1097-1099

Tzahar E, Levkowitz G, Karunagaran D, Yi L, Peles E, Lavi S, Chang D, Liu N, Yayon A, Wen D, et al. ErbB-3 and ErbB-4 function as the respective low and high affinity receptors of all Neu differentiation factor/heregulin isoforms. J Biol Chem 1994, **269**:25226-25233

Urbanski SJ, Edwards DR, Maitland A, Leco KJ, Watson A, Kossakowska AE. Expression of metalloproteinases and their inhibitors in primary pulmonary carcinomas. Br J Cancer 1992, **66**:1188-1194

Vachtenheim J, Horakova I, Novotna H, Opaalka P, Roubkova H. Mutations of K-ras oncogene and absence of H-ras mutations in squamous cell carcinomas of the lung. Clin Cancer Res 1995, **1**:359-365

Vasicek TJ, Zeng L, Guan XJ, Zhang T, Costantini F, Tilghman SM. Two dominant mutations in the mouse fused gene are the result of transposon insertions. Genetics 1997, **147**:777-786

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science 1995, **270**:484-487

Velculescu VE, El-Deiry WS. Biological and clinical importance of the p53 tumor suppressor gene. Clin Chem 1996, **42**:858-868

Verbeek S, Izon D, Hofhuis F, Robanus-Maandag E, te Riele H, van de Wetering M, Oosterwegel M, Wilson A, MacDonald HR, Clevers H. An HMG-box-containing T-cell factor required for thymocyte differentiation. Nature 1995, **374**:70-74

Vernos I, Karsenti E. Motors involved in spindle assembly and chromosome segregation. Curr Opin Cell Biol 1996, **8**:4-9

Vleminckx K, Vakaet L Jr, Mareel M, Fiers W, van Roy F. Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell 1991, **66**:107-119

Vleminckx K, Wong E, Guger K, Rubinfeld B, Polakis P, Gumbiner BM. Adenomatous polyposis coli tumor suppressor protein has signaling activity in Xenopus laevis embryos resulting in the induction of an ectopic dorsoanterior axis. J Cell Biol 1997, **136**:411-420

Vleminckx K, Kemler R. Cadherins and tissue formation: integrating adhesion and signaling. Bioessays 1999, **21**:211-220

Volm M, Efferth T, Mattern J. Oncoprotein (c-myc, c-erbB1, c-erbB2, c-fos) and suppressor gene product (p53) expression in squamous cell carcinomas of the lung. Clinical and biological correlations. Anticancer Res 1992, **12**:11-20

Vonlanthen S, Heighway J, Tschan M, Borner MM, Altermatt HJ, Kappeler A, Tobler A, Fey M, Thatcher N, Yarbough WG, Betticher D. Expression of p16INK4a/p16, and p19ARF/p16β is frequently altered in non-small cell lung cancer and correlates with p53 overexpression. Oncogene 1998, **17**:2779-2785

Wadsworth P. Regional regulation of microtubule dynamics in polarized, motile cells. Cell Motil Cytoskeleton 1999, **42**:48-59

Wagner KU, Wall RJ, St-Onge L, Gruss P, Wynshaw-Boris A, Garrett L, Li M, Furth PA, Hennighausen L. Cre-mediated gene deletion in the mammary gland. Nucleic Acids Res 1997,**25**:4323-4330

Wainwright BJ, Scambler PJ, Stanier P, Watson EK, Bell G, Wicking C, Estivill X, Courtney M, Boue A, Pedersen PS, et al. Isolation of a human gene with protein sequence similarity to human and murine int-1 and the Drosophila segment polarity mutant wingles EMBO J 1988, 7:1743-1748

Wallis YL, Macdonald F, Hulten M, Morton JE, McKeown CM, Neoptolemos JP, Keighley M, Morton DG. Genotype-phenotype correlation between position of constitutional APC gene mutation and CHRPE expression in familial adenomatous polyposis. Hum Genet 1994, **94**:543-548

Waltzer L, Bienz M. Drosophila CBP represses the transcription factor TCF to antagonize Wingless signalling. Nature 1998, **395**:521-525

Walz W, Wuttke WA. Resistance of astrocyte electrical membrane properties to acidosis changes in the presence of lactate. Brain Res 1989, **504**:82-86

Wang J, Shackleford GM. Murine Wnt10a and Wnt10b: cloning and expression in developing limbs, face and skin of embryos and in adults. Oncogene 1996a, **13**:1537-1544

Wang Y, Krushel LA, Edelman GM. Targeted DNA recombination in vivo using an adenovirus carrying the cre recombinase gene. Proc Natl Acad Sci *U S A* 1996b **93**:3932-3936

Wang JY. Retinoblastoma protein in growth suppression and death protection. Curr Opin Genet Dev 1997, **7**:39-45

Wang YK, Samos CH, Peoples R, Perez-Jurado LA, Nusse R, Francke U. A novel human homologue of the Drosophila frizzled wnt receptor gene binds wingless protein and is in the Williams syndrome deletion at 7q11.23. Hum Mol Genet 1997b, **6**:465-472

Watanabe M, Kakiuchi H, Kato H, Shiraishi T, Yatani R, Sugimura T, Nagao M. APC gene mutations in human prostate cancer. Jpn J Clin Oncol 1996, **26**:77-81

Waterman ML, Fischer WH, Jones KA. A thymus-specific member of the HMG protein family regulates the human T cell receptor C alpha enhancer. Genes Dev 1991, **5**:656-669

Waterman-Storer CM, Salmon E. Positive feedback interactions between microtubule and actin dynamics during cell motility. Curr Opin Cell Biol 1999, **11**:61-67

Weber JL, May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 1989, **44**:388-396

Weinberg RA. The retinoblastoma protein and cell cycle control. Cell 1991, **81**:323-330

Weiss EE, Kroemker M, Rudiger AH, Jockusch BM, Rudiger M. Vinculin is part of the cadherin-catenin junctional complex: complex formation between alpha-catenin and vinculin. J Cell Biol 1998, **141**:755-764

Weng S, Tsuchiya E, Satah Y, Kitaagawa T, Nakagowa K, Sugano H. Multiple atypical adenomatous hyperplasia of type II pneumocytes and bronchioloalveolar carcinoma. Histopathology 1990, **16**:101-103

Weng SY, Tsuchiya E, Kasuga T, Sugano H. Incidence of atypical bronchioloalveolar cell hyperplasia of the lung:relation to histological subtypes of lung cancer. Virchows Arch 1992, **420**:463-471

Weston A, Caporaso NE, Perrin LS, Sugimura H, Tamai S, Krontiris TG, Trump BF, Hoover RN, Harris CC. Relationship of H-ras-1, L-myc, and p53 polymorphisms with lung cancer risk and prognosis. Enviro Health Perspect 1992, **98**:61-67

Weston A, Godbold JH. Polymorphisms of H-ras-1 and p53 in breast cancer and lung cancer: a meta-analysis. Environ Health Perspect 1997, **105**:919-926

Westra WH, Slebos RJ, Offerhaus GJ, Goodman SN, Evers SG, Kensler TW, Askin FB, Rodenhuis S, Hruban RH. K-ras oncogene activation in lung adenocarcinomas from former smokers. Evidence that K-ras mutations are an early and irreversible event in the development of adenocarcinoma of the lung. Cancer 1993, **72**:432-438

Westra WH, Baas IO, Hruban RH, Askin FB, Wilson K, Offerhaus GJ, Slebos RJ. K-ras oncogene activation in atypical alveolar hyperplasias of the human lung. Cancer Res 1996, **56**:2224-2228

van de Wetering M, Oosterwegel M, Holstege F, Dooyes D, Suijkerbuijk R, Geurts van Kessel A, Clevers H. The human T cell transcription factor-1 gene. Structure, localization, and promoter characterization. J Biol Chem 1992, **267**:8530-8536

van de Wetering M, Cavallo R, Dooijes D, van Beest M, van Es J, Loureiro J, Ypma A, Hursh D, Jones T, Bejsovec A, Peifer M, Mortin M, Clevers H. Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene dTCF. Cell 1997, **88**:789-799

White GR, Santibrainez-Koref M, Heighway J, Thatcher N. Constitutional frequencies of c-Ha-ras alleles in patients with different types of lung cancer. Br J Cancer 1990, **61**:186

White S, Bubb VJ, Wyllie AH. Germline APC mutation (Gln^{1317}) in a cance-Prone family that does not result in familial adenomatous polyposis. Genes Chrom Cancer 1996, **15**:122-128

Wieland I, Bohm M. Frequent allelic deletion at a novel locus on chromosome 5 in human lung cancer. Cancer Res 1994, **54**:1772-1774

Willert K, Brink M, Wodarz A, Varmus H, Nusse R. Casein kinase 2 associates with and phosphorylates dishevelled. EMBO J 1997, **16**:3089-3096

Wilson CL, Heppner KJ, Labosky PA, Hogan BL, Matrisian LM. Intestinal tumorigenesis is suppressed in mice lacking the metalloproteinase matrilysin. Proc Natl Acad Sci U S A 1997, **94**:1402-1407

Wink DA, Kasprzak KS, Maragos CM, Elespuru RK, Misra M, Dunams TM, Cebula TA, Koch WH, Andrews AW, Allen JS, et al. DNA deaminating ability and genotoxicity of nitric oxide and its progenitors. Science 1991, **254**:1001-1003

Wistuba II, Behrens C, Milchgrub S, Bryant D, Hung J, Minna JD, Gazdar AF. Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. Oncogene 1999, **18**:643-650

Wittinghofer A, Nassar N. How Ras-related proteins talk to their effectors. Trends Biochem Sci 1996, **21**:488-491

Witty JP, McDonnell S, Newell KJ, Cannon P, Navre M, Tressler RJ, Matrisian LM. Modulation of matrilysin levels in colon carcinoma cell lines affects tumorigenicity in vivo. Cancer Res 1994, **54**:4805-4812

Wodarz A, Nusse R. Mechanisms of Wnt signaling in development. Annu Rev Cell Dev Biol 1998;**14**:59-88

Wong MH, Hermiston ML, Syder AJ, Gordon JI. Forced expression of the tumor suppressor adenomatosis polyposis coli protein induces disordered cell migration in the intestinal epithelium. Proc Natl Acad Sci U S A 1996, **93**:9588-9593

Woods DF, Bryant PJ. Molecular cloning of the lethal(1)discs large-1 oncogene of Drosophila. Dev Biol 1989, **134**:222-235

Woods DF, Bryant PJ. The discs-large tumor suppressor gene of Drosophila encodes a guanylate kinase homolog localized at septate junctions. Cell 1991, **66**:451-464

World Health Organisation. Histological typing of lung tumours:Second edition. Am J Clin Pathol 1982; **77**:123-136

Xiao S, Li D, Corson JM, Vijg J, Fletcher JA. Codeletion of p15 and p16 genes in primary non-small cell lung carcinoma. Cancer Res 1995, **55**:2968-2971

Xu H, Hu S, Cagle P, Moore GE, Benedict WF. Absence of retinoblastoma protein expression in primary non-small cell lung carcinomas. Cancer Res 1991, **51**:2735-2739

Xu HJ, Quinlan DC, Davidson AG, Hu SX, Summers CL, Li J, Benedict WF. Altered retinoblastoma protein expression and prognosis in early-stage non-small-cell lung carcinoma. J Natl Cancer Inst 1994, **86**:695-699

Xu Y, Guo DF, Davidson M, Inagami T, Carpenter G. Interaction of the adaptor protein Shc and the adhesion molecule cadherin. J Biol Chem 1997, **272**:13463-13466

Yamamoto H, Kishida S, Uochi T, Ikeda S, Koyama S, Asashima M, Kikuchi A. Axil, a member of the Axin family, interacts with both glycogen synthase kinase 3beta and beta-catenin and inhibits axis formation of Xenopus embryos. Mol Cell Biol 1998, **18**:2867-2875

Yamamoto H, Kishida S, Kishida M, Ikeda S, Takada S, Kikuchi A. Phosphorylation of axin, a Wnt signal negative regulator, by glycogen synthase kinase-3beta regulates its stability. J Biol Chem 1999, **274**:10681-10684

Yamashita N, Minamoto T, Ochiai A, Onda M, Esumi H. Frequent and characteristic K-ras activation in aberrant crypt foci of colon. Is there preference among K-ras mutants for malignant progression? Cancer 1995, **75**:1527-1533

Yanagawa S, van Leeuwen F, Wodarz A, Klingensmith J, Nusse R. The dishevelled protein is modified by wingless signaling in Drosophila. Genes Dev 1995, **9**:1087-1097

Yap AS, Brieher WM, Gumbiner BM. Molecular and functional analysis of cadherin-based adherens junctions. Annu Rev Cell Dev Biol 1997, **13**:119-146

Yap AS, Niessen CM, Gumbiner BM. The juxtamembrane region of the cadherin cytoplasmic tail supports lateral clustering, adhesive strengthening, and interaction with p120ctn. J Cell Biol 1998, **141**:779-789

Yarden Y, Schlessinger J. Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor. Biochemistry 1987, **26**:1443-1451

Yashima K, Nakamori S, Murakami Y, Yamaguchi A, Hayashi K, Ishikawa O, Konishi Y, Sekiya T. Mutations of the adenomatous polyposis coli gene in the mutation cluster region: comparison of human pancreatic and colorectal cancers. Int J Cancer 1994, **59**:43-47

Yesner R *et al.* International Histological Classification of Tumours. No.1 : Histological Typing of Lung Tumours. 2nd ed. Geneva, World Health Organisation 1982

Yost C, Torres M, Miller JR, Huang E, Kimelman D, Moon RT. The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. Genes Dev 1996, **10**:1443-1454

You M, Candrian U, Maronpot RR Stoner GD, Anderson MW. Activation of the Ki-ras protooncogene in spontaneously occuring and chemically induced lung tumors of the strain A mouse. Proc Natl Acad Sci USA 1989, **86**:3070-3074

Yu X, Waltzer L, Bienz M. A new Drosophila APC homologue associated with adhesive zones of epithelial cells. Nat Cell Biol 1999, **1**:144-151

Yunis J, Ramsay N. Retinoblastoma and subband deletion of chromosome 13. Am J Dis Child 1978, **132**:161-163

Zeng L, Fagotto F, Zhang T, Hsu W, Vasicek TJ, Perry WL 3rd, Lee JJ, Tilghman SM, Gumbiner BM, Costantini F. The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. Cell 1997, **90**:181-192

Zhang YJ, Jiang W, Chen CJ, Lee CS, Kahn SM, Santella RM, Weinstein IB. Amplification and overexpression of cyclin D1 in human hepatocellular carcinoma. Biochem Biophys Res Commun 1993, **196**:1010-1016

Zhang Y, Riesterer C, Ayrall AM, Sablitzky F, Littlewood TD, Reth M. Inducible site-directed recombination in mouse embryonic stem cells. Nucleic Acids Res 1996, **24**:543-548

Zheng L, Zhang J, Carthew RW. frizzled regulates mirror-symmetric pattern formation in the Drosophila eye. Development 1995, **121**:3045-3055

APPENDIX A: Solutions

General Solutions

TE

10mM Tris (pH 8.0)

1mM EDTA

10 x TBE (tris-borate-EDTA buffer)

0.89 M Tris borate (pH 8.3)

0.02M EDTA

PBS

0.138M NaCl

2.7mM KCl

pH to 7.4 and autoclaved

(x10) gel loading buffer

0.25% bromophenol blue

25% Ficoll

Medium

SOC

2% bactotryptone

0.5% yeast extract

10mM NaCl

2.5mM KCl

10mM $MgCl_2$

20mM glucose

Preperation of 1L

1) Dissolve 20g Tryptone, 5g Yeast Extract and 0.5g NaCl in 950ml deionised water.

2) Prepare a 250mM KCl stock solution

3) Add 10ml of KCl stock solution to the solution in step 1.

3) Adjust pH to 7.0 with 5M NaOH, then bring volume to 980ml with deionised water.

4) Prepare 1M solution of $MgCl_2$

5) Autoclave both solutions

6) Prepar a 2M solution of glucose and filter sterilise.

7) Add 10ml of filter sterilised 2M glucose solution and 10ml $MgCl_2$ solution to cooled autoclaved solution.

LB (Luria-Bertani) medium and plates

1% Bacto-Tryptone

0.5% Bacto-Yeast Extract

1% NaCl

pH7.0

Medium: Preperation of 1L

1) Dissolve 10g Bacto-Tryptone, 5g Bacto-Yeast Extract and 10g NaCl in 950ml deionised water.
2) Adjust pH to 7.0 with NaOH and ajust volume to 1L
3) Autoclave and allow to cool to 55^{o}C and add antibiotic if needed
4) Store at RT or at 4^{o}C.

LB plates

1) Prepare LB medium as above, once cooled to 55^{o}C, add antibiotic (stock solution of ampicilin to a final concentration of 50μg/ml) and pour into 10cm plates.
2) Allow to harden, then invert and store at 4^{o}C

Ampicillin Stock solution 100mg/ml

1) Dissolve 10g of Amicillin into 10ml of sterile deionised water
2) Aliquote into 500ul and store at –20^{o}C

X-gal stock solution 40mg/ml

1) Dissolve 400mg X-gal in 10ml dimethylformamide
2) Protect from light by storing in a brown bottle at –20^{o}C

To add to previously made agar plates, pipette 40μl of stock solution onto each plate, spread evenly, and let dry 15 minutes. Protect plates from light. Store at 4^{o}C

DNA extraction

Lysis Buffer

10mM Tris-HCl pH 8.3

50mM KCl

0.45% (v/v) Tween-20

2.5mM $MgCl_2$

Single Strand Conformational Polymorphism

Denaturing solution

0.5M NaOH

10mM EDTA

Stop Solution

95% formamide

10mM EDTA (pH 8)

0.01% (w/v) bromophenol blue

0.01% (w/v) xylene cyanol

APC Immunohistochemistry

DAB

5mg/ml diaminobenzidine

0.2M Tris-HCl pH 7.6

0.01M imidazole

0.05% (v/v) hydrogen peroxide

Sequencing

(2x) Binding Wash

10mM Tris-HCl (pH 7.5)

1mM EDTA

2.0M NaCl

plasmid reaction buffer

1.0M Tris-HCl pH 7.5

100mM $MgCl_2$

250mM NaCl

Sequencing Stop Solution

95% formamide

10mM EDTA (pH 8)

0.01% (w/v) bromophenol blue

0.01% (w/v) xylene cyanol

Reaction Buffer

200mM Tris-HCl (pH7.5)

100mM $MgCl_2$

250mM NaCl

Labelling Mix

7.5μM dGTP

7.5μM dCTP

7.5μM dTTP

6% Sequencing Gel

6% acrylamide:bisacrylamide (40:1)

8M Urea

1xTBE

1μl 25% ammonium persulphate per ml of gel

1μl TEMED per ml gel

Ligations

(5x) Reaction Buffer

250mM Tris-HCl (pH 7.6)

50mM $MgCl_2$

5mM dATP

5mM dithiothreitol

25% polyethylene glycol-8000

10 X Ligation Buffer

60Mm Tris-HCL, pH7.5

60mM MgCl2

50mM NaCl

1mg/ml bovine serum albumin

70mMβ-mercaptoethanol

1mM ATP

20mM dithiothreitol

10mMspermidine

Wizard Mini Prep Solutions

Cell resuspension buffer

50mM Tris-HCl (pH 7.5)

10mM EDTA

100μg/ml RNase A

Cell Lysis Solution

0.2M NaOH

35mM SDS

Neutralising Solution

1.32M potassium acetate, pH 4.8

Column Wash

80mM potassium acetate

8.3mM Tris-HCl pH 7.5

40μM EDTA

55% (v/v) ethanol

Southern Blotting

(20x) SSC

3M NaCl

0.3 M Sodium Citrate pH 7.0

Pre-Hybridisation solution

6 X SSC

5 X Denhardt's

17mM SDS

Denhardts Reagent (100x)

2% (w/v) Ficoll

2% (w/v) Polyvinyl pyrollidone

2% (w/v) Bovine serum albumin (Fraction V)

High Prime Mix

4 units of Klenow polymerase

0.125 mM of each of dATP, dGTP, dTTP

5x reaction buffer in 50% glycerol

Stringency wash 1

2x SSC

0.1% SDS

Stringency wash 2

1xSSC

0.1% SDS

Stringency wash 3

0.1xSSC

0.1% SDS

RNA

DEPC treated water

0.01% (w/v) diethylpyrocarbonate (DEPC) in double distilled water

Stand overnight at room temperature

Autoclave

10x DNase reaction buffer

200mM Tris-HCl (pH 8.4)

20mM $MgCl_2$

500mM KCl

Embryonic Fibroblasts

Embryonic fibroblast Culture medium

MEM Glasgow BHK-21 (Life technologies, UK)

5% fetal calf serum

5% neonatal calf serum

1mM MEM non-essential amino acids (Life Technologies)

1mM of sodium pyrutate (Gibco, UK)

Embryonic Fibroblast freezing medium

25% fetal calf serum

25% normal calf serum

40% MEM Glasgow BHK-21 medium

10% DMSO

Cell fixative

0.05% gluteraldehyde in PBS

25% stock solution was stored at -20°C and diluted 1/500 in PBS prior to use.

β-galactosidase stain

20mM NaH_2PO_4

80mM Na_2HPO_4

1.3mM $K_3Fe(CN)_6$

3mM $K_4Fe(CN)_6$

1mg/ml x-gal.

Appendix B: Suppliers

Amersham Life Science Ltd. Amersham Place Little Chalfont Buckinghamshire HP7 9NA	BDH Merck Ltd. Hunter Boulevard Magna Park Lutterworth Leicester
Bio-Rad Laboratories Ltd. Bio-Rad House Maylands Avenue Hemel Hempstead Hertfordshire	Boehringer Manheim UK (Diagnostic and Biochemicals) Limited. Bell Lane Lewes East Sussex
Dako Ltd 22 The Arcade The Octagon High Wycombe Buckinghamshire HP11 2HT	Dynal UK Ltd Station House 26 Grove Street New Ferry Wirral Merseyside
Flowgen Lynn Lane Shenstone Staffordshire WS14 0EE	Gibco (BRL). Life Technologies Ltd. 3 Fountain Drive Paisley PA4 9RF
Hybaid Ltd 111-113 Waldegrave Road Teddington Middlessex TW11 8LL	Hoefer Ltd Croft Road Newcastle-Under-Lyme Staffordshire ST5 0TW

Invitrogen BV PO Box 2312,9704 CH Groningen The Netherlands	Kodak Ltd IBI Ltd 36 Clifton Road Cambridge CB1 4ZR
New England Biolabs (UK) Ltd. Knowl Piece Wilbury Way Hitchin	Oncogene Research Products 84 Rogers Street Cambridge MA 02142
Oswell DNA Service Dept. of Chemistry University of Edinburgh Kings Buildings West Mains Road Edinburgh	Pharmacia Biotech 23 Grosvenor Road St Albans Herts AL1 3AW
Promega UK Ltd Delta House, Chilworth Research Centre, Southampton. SO16 7NS	Quiagen Ltd Unit 1Tillingbourne Court Station Road Dorking Surrey RH4 1HJ
Sigma Ltd Fancy Road Poole Dorset BH17 7NH	Stratagene Europe Gebouw California Hogehilweg 15 1101 CB Amsterdam Zuidoost The Netherlands

APPENDIX C: NSCLC database; analysis of LOH at 5q21

Abbreviations

B.ADC- bronchial adenocarcinoma P. ADC- parenchymal adenocarcinoma

SCC-squamous cell carcinoma

NI- non-informative IRET- Informative retained

ILOH- informative loss of heterozygosity

A1/a2- allele1/allele 2

CASE	UB NO.	AGE	SEX	TNM	HISTOLOGY	APC 3'	APC EX.11	MCC EX.10	MCC 3'
1	84/00950	65	F	T2N2	B. ADC	NI	IRET	NI	IRET
2	84/01170	83	M	T1N0	P. ADC	IRET	NT	NI	IRET
3	84/01687	65	F	T2N0	ADC	NT	IRET	NI	IRET
4	84/01742	66	M	T2N0	ADC	ILOHa1	ILOHa2	NI	NI
5	84/03245	68	M	T1N0	P. ADC	ILOHa2	NT	ILOHa1	NT
6	84/03801	54	M	T2N1	P. ADC	NT	IRET	NT	IRET
7	84/03921	50	F	T2N0	ADC	NT	IRET	NI	IRET
8	84/09939		F	T2N0	B. ADC	ILOHa2	NI	ILOHa1	NT
9	84/10202	58	F	T2N1	P. ADC	NI	NI	IRET	NI
10	84/14695	65	F	T3N1	ADC	ILOHa1	ILOHa2	NI	ILOHa2
11	84/16024	50	F	T2N2	ADC	NI	NI	NI	IRET
12	84/16468	71	M	T2N0	ADC	NI	NI	ILOHa1	ILOHa1
13	85/05898	62	F	T2N1	P. ADC	NT	NT	ILOHa2	ILOHa1
14	85/07275	69	F	T2N0	B. ADC	NT	ILOHa1	ILOHa2	NI
15	85/17016	62	F	T1N0	T1 P. ADC	T1 IRET	NT	T1 IRET	NT
				T2N0	T2 adenosquame	T2 ILOHa2		T2 ILOHa1	
16	89/25140	69	F	T2N2	SCC	IRET	IRET	IRET	NT
17	89/25506	57	M	T2N2	ADC	IRET	NI	NI	NI
18	89/26713	68	M	T2N0	P. ADC	IRET	IRET	IRET	IRET
19	89/27480	60	M	T2N0	P. ADC	IRET	IRET	NI	IRET
20	89/28300	56	M	T2N1	SCC	ILOHa2	ILOHa1	NI	NI
21	89/29807	45	F	T3N0	ADC	IRET	NI	NT	IRET
22	90/00047	59	M	T2N0	SCC	ILOHa1	ILOHa2	ILOHa2	NT
23	90/00413	46	M	T1N0	P. ADC	NI	NI	NI	NI

24	90/00417	69	M	T2N0	SCC	NI	NI	NI	IRET
25	90/00690	57	M	T3N0	ADC	NI	NI	NI	IRET
26	90/00724	63	F	T2N0	B.ADC	NT	IRET	IRET	NT
27	90/01327	58	M	T2N0	ADC	IRET	IRET	NI	NI
28	90/02125	56	M	T2N0	SCC	NI	NI	NI	NI
29	90/02162	70	F	T2N1	ADC	NT	ILOHa2	ILOHa2	ILOHa2
30	90/02257	62	M	T2N1	SCC	NT	IRET	IRET	NT
31	90/02501	76	M	T2N0	SCC	IRET	IRET	IRET	IRET
32	90/02736	75	M	T2N1	P. ADC	NT	IRET	IRET	NT
33	90/02956	72	M	T3N0	SCC	NT	IRETx2	NIx2	IRETX2
34	90/03496	66	M	T2N2	SCC	NI	ILOHa2	NI	ILOHa1
35	90/03498	63	M	T3N2	T1 P. ADC T2 SCC	IRET X2	IRETX2	IRETX2	IRETX2
36	90/03739	68	M	T2N1	SCC	NI	NI	NI	NI
37	90/03971	75	M	T2N1	P. ADC	NT	ILOHa2	NI	ILOHa1
38	90/04074	74	M	T2N1	SCC	ILOHa1	ILOHa2	NI	NI
39	90/04564	53	M	T2N0	P. ADC	NI	NI	NI	IRET
40	90/04759	70	M	T2N0	SCC	ILOHa1	NI	ILOHa2	ILOHa2
41	90/05104	63	M		ADC	NIx2	T1 ILOHa2 T2 RET	NIx2	NI
42	90/07250	66	M	T2N0	ADC	IRET	IRET	IRET	NT
43	90/07387	57	M	T3N2	ADC	IRET	NI	IRET	NT
44	90/08639	64	M	T2N0	ADC	ILOHa1	ILOHa2	NI	NI
45	90/09521	69	M	T3N0	ADC	ILOHa2	NI	ILOHa1	NT
46	90/10230	51	F	T2N0	P. ADC	ILOHa2	NI	ILOHa2	NT
47	90/11758	69	M	T2N0	B.ADC	IRET	NI	NI	NI
48	90/14151	64	M	T2N0	ADC	NT	NI	NI	NI
49	90/14771	55	M	T2N0	B. ADC	NT	IRET	NI	IRET
50	90/17878	62	F	T1N0	ADC	IRET	NI	IRET	NT
51	90/17920	66	M	T1N1	SCC	NI	IRET	IRET	NI
52	90/18877	70	M	T2N1	P. ADC	IRET	NT	NI	IRET
53	90/21957	66	F	T2N1	P. ADC	IRET	IRET	IRET	NT
54	90/22073	73	F	T2N0	ADC	NI	NI	IRET	IRET
55	90/22081	66	F	T2N0	ADC	NI	NI	IRET	NI
56	90/22482	51	M	T3N2	P. ADC	ILOHa2	ILOHa2	NI	ILOHa2
57	90/22711	49	F	T1N0	P. ADC	NT	IRET	IRET	NT
58	90/23269	66	F	T2N2	ADC	NI	ILOHa2	NI	NI
59	90/23653	67	M	T2N0	B.ADC	IRET	NT	NI	IRET
60	90/24563	69	M	T3N2	ADC	ILOHa2	NT	ILOHa2	NT
61	90/25526	53	M	T2N1	ADC	IRET	IRET	NI	NI
62	90/257941	71	M	T2N0	ADC	ILOHa1	ILOHa2	ILOHa2	NT
63	90/26831	64	F	T3N2	ADC	NI	NI	IRET	NI
64	90/27759	74	M	T2N0	P. ADC	NI	IRET	NI	NI

65	90/27957	71	M	T2N1	SCC	IRET	NI	NI	IRET
66	90/27988	63	M	T2N0	SCC	NI	NI	NI	NT
67	90/28125	63	M	T2N0	SCC	NI	NI	NI	ILOHa1
68	91/00297	63	M	T2N0	SCC	NI	IRET	NI	IRET
69	91/00486	82	M	T2N1	SCC	NI	NI	ILOHa1	ILOHa1
70	91/01945	75	M	T1N0	P. ADC	IRET	IRET	IRET	IRET
71	91/02175	54	F	T2N2	ADC	NI	NI	NI	NI
72	91/02882	73	M	T2N0	SCC	NI	NI	ILOHa2	ILOHa2
73	91/04010	79	M	T2N0	B. ADC	ILOHa1	ILOH	NI	NI
74	91/05849	67	M	T1N0	SCC	NI	IRET	NI	NI
75	91/06104	37	M	T2N0	SCC	NI	NI	NI	NI
76	91/06385	64	M	T3N0	SCC	NI	NI	IRET	NI
77	91/07148	59	M	T2N0	P. ADC	IRET	IRET	NI	IRET
78	91/0665	61	F	T2N0	SCC	IRET	IRET	NI	NI
79	91/07238	55	M	T2N0	P. ADC	NI	NI	ILOHa1	NI
80	91/07877	66	M	T2N0	SCC	NI	NI	ILOHa1	ILOHa1
81	91/08090	51	M	T2N2	SCC	ILOHa1	ILOHa2	NI	NI
82	91/08381	57	F	T3N2	SCC	ILOHa1	ILOHa2	NI	ILOHa2
83	91/08444	63	M	T2N0	ADC	IRET	IRET	NT	NT
84	91/08748	46	M	T2N1	SCC	IRET	NI	NI	NI
85	91/08925	67	M	T1N1	SCC	IRET	IRET	IRET	IRET
86	91/09054	59	M	T2N0	P. ADC	IRET	IRET	NI	NI
87	91/09159	60	M	T2N2	SCC	IRET	NI	IRET	NI
88	91/09363	84	M	T2N1	P. ADC	ILOHa1	NI	NI	ILOHa1
89	91/09369	73	M	T2N0	SCC	ILOHa2	NI	ILOHa1	NT
90	91/09534	76	F	T2N0	SCC	IRET	IRET	IRET	NI
91	91/09940	79	M	T2N0	SCC	IRET	IRET	NI	NI
92	91/10095	63	M	T3N1	P. ADC	IRET	IRET	IRET	NT
93	94/00111	72	F	T2N0	B. ADC	NI	NI	ILOHa1	NI
94	94/01642	58	F	T3N0/ T1N0	ADC x2	IRET	NI	NIx2	IRET
95	94/02910	64	M	T2N1	B. ADC	NI	IRET	NI	NI
96	94/04401	66	M	T2N0	B. ADC	ILOHa2	ILOHa1	NI	NI
97	94/04621	60	M	T2N0	B. ADC	NI	ILOHa2	NI	ILOHa2

Appendix D: NSCLC database; analysis of *K-RAS* muations in ADC

Abbreviations

B.ADC- bronchial adenocarcinoma P. ADC- parenchymal adenocarcinoma

CASE	UB NO.	AGE	SEX	TNM	HISTOLOGY	K-RAS codon 12	Mutation
1	84/00950	65	F	T2N2	B. ADC	-	
2	84/01170	83	M	T1N0	P. ADC	-	
3	84/01687	65	F	T2N0	ADC	-	
4	84/01742	66	M	T2N0	ADC	-	
5	84/03245	68	M	T1N0	P. ADC	-	
6	84/03801	54	M	T2N1	P. ADC	-	
7	84/03921	50	F	T2N0	ADC	-	
8	84/09939		F	T2N0	B. ADC	-	
9	84/10202	58	F	T2N1	P. ADC	-	
10	84/14695	65	F	T3N1	ADC	-	
11	84/16024	50	F	T2N2	ADC	+	G-T pos 2
12	84/16468	71	M	T2N0	ADC	+	G-T pos.1
13	85/05898	62	F	T2N1	P. ADC	-	
14	85/07275	69	F	T2N0	B. ADC	-	
15	85/17016	62	F	T1N0	T1 P. ADC	T1 +	G-T pos.1
				T2N0	T2 adenosquame	T2 -	
17	89/25506	57	M	T2N2	ADC	-	
18	89/26713	68	M	T2N0	P. ADC	+	G-T pos.1
19	89/27480	60	M	T2N0	P. ADC	+	G-T pos.2
21	89/29807	45	F	T3N0	ADC	-	
23	90/00413	46	M	T1N0	P. ADC	+	G-T pos. 2
25	90/00690	57	M	T3N0	ADC	-	
26	90/00724	63	F	T2N0	B.ADC	-	
27	90/01327	58	M	T2N0	ADC	-	
29	90/02162	70	F	T2N1	ADC	+	G-T pos 1
32	90/02736	75	M	T2N1	P. ADC	-	
35	90/03498	63	M	T3N2	P. ADC	+	G-T pos. 1
37	90/03971	75	M	T2N1	P. ADC	-	
39	90/04564	53	M	T2N0	P. ADC	+	G-T pos.1
41	90/05104	63	M		ADC	-	
42	90/07250	66	M	T2N0	ADC	+	G-T pos.1
43	90/07387	57	M	T3N2	ADC	-	
44	90/08639	64	M	T2N0	ADC	-	

45	90/09521	69	M	T3N0	ADC	-	
46	90/10230	51	F	T2N0	P. ADC	-	
47	90/11758	69	M	T2N0	B.ADC	-	
48	90/14151	64	M	T2N0	ADC	-	
49	90/14771	55	M	T2N0	B.ADC	-	
50	90/17878	62	F	T1N0	ADC	-	
52	90/18877	70	M	T2N1	P. ADC	-	
53	90/21957	66	F	T2N1	P. ADC	-	
54	90/22073	73	F	T2N0	ADC	-	
55	90/22081	66	F	T2N0	ADC	+	G-T pos.2
56	90/22482	51	M	T3N2	P. ADC	-	
57	90/22711	49	F	T1N0	P. ADC	+	G-T pos. 1
58	90/23269	66	F	T2N2	ADC	-	
59	90/23653	67	M	T2N0	B.ADC	-	
60	90/24563	69	M	T3N2	ADC	-	
61	90/25526	53	M	T2N1	ADC	-	
62	90/257941	71	M	T2N0	ADC	-	
63	90/26831	64	F	T3N2	ADC	-	
64	90/27759	74	M	T2N0	P. ADC	-	
70	91/01945	75	M	T1N0	P. ADC	-	
71	91/02175	54	F	T2N2	ADC	-	
73	91/04010	79	M	T2N0	B. ADC	-	
77	91/07148	59	M	T2N0	P. ADC	+	G-A pos.2
79	91/07238	55	M	T2N0	P. ADC	+	G-T pos.2
83	91/08444	63	M	T2N0	ADC	-	
86	91/09054	59	M	T2N0	P. ADC	-	
88	91/09363	84	M	T2N1	P. ADC	+	G-T pos. 1
92	91/10095	63	M	T3N1	P. ADC	-	
93	94/00111	72	F	T2N0	B. ADC	-	
94	94/01642	58	F	T3N0/T1 N0	ADC x2	T1- T2 +	G-T pos.1
95	94/02910	64	M	T2N1	B. ADC	-	
96	94/04401	66	M	T2N0	B. ADC	-	
97	94/04621	60	M	T2N0	B. ADC	-	

Appendix E: Database for murine breeding program

Abbreviations: WT- wildtype; Tg- Transgene; D.O.B- date of birth ; NEG –negative; MIN (genotype $Apc^{+/-}$)

No.	Sex	Apc	Tg	Mating	D.O.B	Comments
Founder CC	F	WT	POSITIVE	Micro-injection		
1	M	WT	NEG	CC X MIN	10.03.98	
2	M	WT	POSITIVE	CC X MIN	10.03.98	
3	M	WT	NEG	CC X MIN	30.3.98	
4	M	WT	POSITIVE	CC X MIN	30.3.98	Cull for Tg expression experiments
5	F	WT	NEG	CC X MIN	30.3.98	
6	M	WT	POSITIVE	CC X MIN	30.3.98	Cull for Tg expression experiments
7	F	WT	NEG	CC X MIN	30.3.98	
8	M	WT	NEG	CC X MIN	30.3.98	
9	M	WT	NEG	CC X MIN	30.3.98	
10	F	HET	POSITIVE	CC X MIN	30.3.98	Cull 26.2.99 ill/colonic adenomas
11	M	WT	POSITIVE	CC X MIN	30.3.98	Cull for Tg expression experiments
12	M	WT	NEG	CC X MIN	30.3.98	
13	F	WT	NEG	2 X MIN	14.05.98	Found dead 16.06.98: Runt
14	M	WT	NEG	2 X MIN	14.05.98	
15	F	WT	NEG	2 X MIN	14.05.98	
16	M	WT	POSITIVE	2 X MIN	19.05.98	
17	M	HET	NEG	2 X MIN	19.05.98	
18	M	HET	POSITIVE	2 X MIN	19.05.98	Cull 26.2.99 ill/colonic adenomas
19	M	HET	POSITIVE	2 X MIN	19.05.98	
20	F	WT	POSITIVE	2 X MIN	01.06.98	Breed with Mgb-c 2 20.07.98 for Efs
21	M	WT	NEG	2 X MIN	01.06.98	
22	M	HET	NEG	2 X MIN	01.06.98	
23	F	WT	NEG	2 X MIN	01.06.98	
24	F	HET	POSITIVE	2 X MIN	01.06.98	
25	F	WT	NEG	2 X MIN	01.06.98	Found dead 06.07.98 (no adenomas)
26	M	HET	NEG	CC X MIN	05.06.98	
27	M	WT	NEG	CC X MIN	05.06.98	
28	F	WT	NEG	CC X MIN	05.06.98	Found dead 06.07.98 (no adenomas)
29	F	WT	NEG	CC X MIN	05.06.98	Found dead 30.6.98 (no adenomas)
30	F	WT	NEG	CC X MIN	05.06.98	

31	F	WT	NEG	CC X MIN	05.06.98	
32	M	WT	NEG	CC X MIN	05.06.98	
33	M	WT	NEG	CC X MIN	05.06.98	
34	F	WT	NEG	CC X MIN	05.06.98	
35	M	WT	NEG	CC X MIN	05.06.98	
36	F	WT	NEG	CC X MIN	05.06.98	Found dead 06.07.98 (no adenomas)
37	F	WT	NEG	11 X MIN (X3)	06.06.98	
38	M	WT	NEG	11 X MIN (X3)	06.06.98	
39	F	WT	NEG	11 X MIN (X3)	06.06.98	
40	M	WT	NEG	11 X MIN (X3)	06.06.98	
41	F	WT	NEG	11 X MIN (X3)	06.06.98	
42	F	WT	NEG	11 X MIN (X3)	06.06.98	
43	M	WT	NEG	4 X MIN (X3)	08.06.98	
44	M	WT	NEG	4 X MIN (x3)	08.06.98	
45	M	WT	NEG	4 X MIN (x3)	08.06.98	
46	M	WT	NEG	4 X MIN (x3)	08.06.98	
47	F	WT	NEG	4 X MIN (x3)	08.06.98	
48	F	WT	NEG	4 X MIN (x3)	08.06.98	
49	M	WT	NEG	4 X MIN (x3)	08.06.98	
50	M	WT	NEG	4 X MIN (x3)	08.06.98	
51	F	WT	NEG	4 X MIN (x3)	08.06.98	
52	M	WT	POSITIVE	4 X MIN (x3)	15.06.98	
53	M	WT	NEG	4 X MIN (x3)	15.06.98	
54	M	HET	POSITIVE	11 X MIN (x3)	12.06.98	
55	M	HET	POSITIVE	11 X MIN (x3)	12.06.98	Cull 02.2.99 ill/colonic adenomas
56	M	HET	NEG	11 X MIN (x3))	12.06.98	Cull 19.01.99 ill/colonic adenomas
57	F	WT	NEG	11 X MIN (x3)	12.06.98	
58	F	HET	NEG	11 X MIN (x3)	12.06.98	
59	M	HET	POSITIVE	11 X MIN (x3)	12.06.98	
60	F	HET	POSITIVE	11 X MIN (x3)	12.06.98	
61	F	WT	NEG	11 X MIN (x3)	12.06.98	
62	F	HET	POSITIVE	11 X MIN (X3)	12.06.98	
63	F	HET	NEG	2 X MIN	10.06.98	
64	M	WT	NEG	2 X MIN	10.06.98	
65	M	HET	NEG	2 X MIN	10.06.98	
66	F	WT	POSITIVE	2 X MIN	10.06.98	
67	F	HET	POSITIVE	2 X MIN	10.06.98	Cull 02.2.99 ill/colonic adenomas

68	F	WT	POSITIVE	2 X MIN	10.06.98	
69	F	WT	POSITIVE	2 X MIN	10.06.98	
70	F	WT	POSITIVE	2 X MIN	10.06.98	
71	F	WT	POSITIVE	2 X MIN	10.06.98	
72	F	WT	NEG	2 X MIN	10.06.98	
73	M	WT	NEG	2 X MIN	10.06.98	
74	M	WT	NEG	10 X MIN	29.06.98	
75	F	HET	NEG	10 X MIN	29.06.98	
76	F	WT	NEG	10 X MIN	29.06.98	
77	M	HET	NEG	10 X MIN	29.06.98	
78	F	HET	POSITIVE	10 X MIN	29.06.98	Cull 02.2.99 ill/colonic adenomas
79	M	HET	NEG	11 X MIN (X3)	03.07.98	
80	M	WT	NEG	11 X MIN (X3)	03.07.98	
81	F	HET	NEG	11 X MIN (X3)	03.07.98	Cull 26/2/99: Mammary tumours/ colonic adenomas
82	M	HET	POSITIVE	11 X MIN (X3)	03.07.98	
83	F	HET	POSITIVE	11 X MIN (X3)	03.07.98	
84	F	HET	NEG	4 X MIN (X3)	9/07/98	
85	F	WT	NEG	4 X MIN (X3)	09.07.98	
86	F	HET	NEG	4 X MIN (x3)	09.07.98	
87	M	HET	POSITIVE	4 X MIN (x3)	09.07.98	
88	M	HET	NEG	4 X MIN (X3)	09.07.98	
89	M	WT	NEG	4 X MIN (X3)	09.07.98	
90	M	WT	POSITIVE	4 X MIN (X3)	09.07.98	
91	F	WT	POSITIVE	4 X MIN (X3)	09.07.98	
92	M	WT	POSITIVE	11 X MIN (X3)	14.07.98	
93	F	WT	NEG	11 X MIN X3)	14.07.98	
94	M	HET	NEG	11 X MIN (X3)	14.07.98	
95	M	HET	NEG	11 X MIN (X3)	14.07.98	
96	F	WT	NEG	11 X MIN (X3)	14.07.98	
97	F	WT	POSITIVE	11 X MIN(X3)	14.07.98	
98	M	HET	POSITIVE	11 X MIN (X3)	14.07.98	
99	M	WT	NEG	11 X MIN (X3)	14.07.98	
100	M	HET	NEG	11 X MIN (X3)	14.07.98	
101	M	WT	NEG	10 X 18	19.08.98	
102	F	HET	POSITIVE	10 X 18	19.08.99	Mate with 18 for embryos
103	M	WT	NEG	10 X 18	19.08.99	

104	F	HET	NEG	10 X 18	19.08.98	
105	F	HET	POSITIVE	10 X 18	19.08.98	Mate with 18 for embryos
106	F	HET	POSITIVE	10 X 18	19.08.98	Mate with 18 for embryos
107	M	HET	NEG	55 X 62	26.08.98	
108	M	HET	POSITIVE	55 X 62	26.08.98	
109	M	HET	NEG	55 X 62	26.08.98	
110	F	WT	POSITIVE	55 X 62	26.08.98	
111	F	HET	POSITIVE	55 X 62	26.08.98	Mate with 18 for embryos
112	M	WT	NEG	59 X 67	02.09.98	
113	M	WT	NEG	59 X 67	02.09.98	
114	M	HET	POSITIVE	59 X 67	02.09.98	
115	M	WT	POSITIVE	59 X 67	02.09.98	
116	M	HET	POSITIVE	59 X 67	02.09.98	
117	M	HET	POSITIVE	59 X 67	02.09.98	
118	F	HET	POSITIVE	59 X 67	02.09.98	
119	F	HET	POSITIVE	54 X 60	07.09.98	
120	M	HET	POSITIVE	54 X 60	07.09.98	
121	M	HET	POSITIVE	54 X 60	07.09.98	
122	F	HET	POSITIVE	54 X 60	07.09.98	
123	F	HET	NEG	54 X 60	07.09.98	
124	F	HET	POSITIVE	19 X 24	07.09.98	
125	F	WT	POSITIVE	19 X 24	07.09.98	
127	F	WT	POSITIVE	19 X 24	07.09.98	
127	F	WT	NEG	19 X 24	07.09.98	
128	F	WT	POSITIVE	10 X 18	11.09.98	
129	M	WT	POSITIVE	10 X 18	11.09.98	
130	F	HET	NEG	10 X 18	11.09.98	
131	M	WT	POSITIVE	10 X 18	11.09.98	
132	M	WT	POSITIVE	10 X 18	11.09.98	
133	F	HET	POSITIVE	10 X 18	11.09.98	
134	M	WT	POSITIVE	10 X 18	11.09.98	
135	F	HET	POSITIVE	55 X 62	20.09.98	
136	F	HET	POSITIVE	55 X 62	20.09.98	
137	M	WT	POSITIVE	55 X 62	20.09.98	
138	F	HET	POSITIVE	55 X 62	20.09.98	
139	F	HET	POSITIVE	55 X 62	20.09.98	
140	M	HET	POSITIVE	55 X 62	20.09.98	

141	F	HET	POSITIVE	55 X 62	20.09.98	Found dead 20.10.98: Runt
142	F	WT	POSITIVE	55 X 62	20.09.98	
143	M	HET	POSITIVE	91 X MIN	23.09.98	
144	M	WT	POSITIVE	91 X MIN	23.09.98	Found dead 20.10.98 mother not feeding
145	F	WT	POSITIVE	91 X MIN	23.09.98	Found dead 20.10.98 mother not feeding
146	F	WT	POSITIVE	91 X MIN	23.09.98	Found dead 20.10.98 mother not feeding
147	F	WT	POSITIVE	91 X MIN	23.09.98	
148	M	WT	NEG	91 X MIN	23.09.98	
149	M	HET	POSITIVE	91 X MIN	23.09.98	
150	M	HET	POSITIVE	54 X 60	01.10.98	
151	M	WT	POSITIVE	54 X 60	01.10.98	
152	M	HET	POSITIVE	54 X 60	01.10.98	
153	M	HET	POSITIVE	54 X 60	01.10.98	
154	M	HET	POSITIVE	59 X 67	02.10.98	
155	F	HET	POSITIVE	59 X 67	02.10.98	

Publications

JOURNAL OF PATHOLOGY, VOL. 180: 33–37 (1996)

LOSS OF HETEROZYGOSITY AT 5q21 IN NON-SMALL CELL LUNG CANCER: A FREQUENT EVENT BUT WITHOUT EVIDENCE OF *APC* MUTATION

CINDY A. COOPER, VIVIEN J. BUBB, NICOLA SMITHSON, ROBERT L. CARTER, SARAH GLEDHILL, DAVID LAMB, ANDREW H. WYLLIE AND FRANCIS A. CAREY*

*Sir Alastair Currie CRC Laboratories, Molecular Medicine Centre, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, U.K.; *Department of Pathology, Ninewells Hospital and Medical School, Dundee DD1 9SY, U.K.*

SUMMARY

Four genetic polymorphisms in the *APC* and *MCC* genes at chromosome 5q21 were analysed for loss of heterozygosity (LOH) in 97 primary squamous carcinomas and adenocarcinomas of the lung. LOH was identified in at least two polymorphic loci in 41 percent of informative cases. There was no significant difference in the frequency of LOH between squamous carcinomas and adenocarcinomas. Within the adenocarcinoma group, however, LOH appeared to be more common in tumours having a bronchial origin (5/9; 56 per cent) than in parenchymal adenocarcinoma (6/21; 29 per cent). All 32 tumours showing LOH at one or more polymorphic sites were examined for mutations in the mutation cluster region (MCR) of *APC* by single-strand conformational polymorphism (SSCP) analysis. Mutations were not detected in any of these cases. We therefore propose that it is likely that a tumour suppressor gene on 5q other than *APC* is involved in the pathogenesis of lung cancer.

KEY WORDS—**non-small cell lung cancer; adenocarcinoma; squamous carcinoma; *APC*; *MCC***

INTRODUCTION

Lung cancer is characterized by a greater heterogeneity of morphological appearances than the other common visceral malignancies.[1] Eight histological classes of primary pulmonary malignancy are described in the World Health Organisation (WHO) classification of lung tumours.[2] From the therapeutic and prognostic viewpoint, the prime distinction to be made is between small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC), with the former having a poorer clinical outlook. A variety of molecular lesions have been described in lung cancer, some of which show correlation with histological classification. Thus, amplification of c-*myc* is often associated with SCLC, while K-*ras* mutations and increased expression of c-*erb*B-2 are seen in NSCLC, particularly adenocarcinoma.[3,4]

Recently there have been a number of reports of genetic losses at chromosome 5q in lung carcinomas.[5–8] Most attention has been paid to SCLC, in which such losses are described in about 80 per cent of cases, but other groups, including our own, have shown losses at 5q in NSCLC.[5,8] Consistent loss of genetic material at a particular chromosomal arm in a neoplasm is evidence for the presence of a tumour suppressor gene at that site. Although other loci on 5q may be implicated, 5q21 is a prime candidate region, as it contains the tumour suppressor gene *APC* (adenomatous polyposis coli) and the candidate tumour suppressor gene *MCC* (mutated in colorectal cancer). *APC* in particular is an important and relatively well-characterized tumour suppressor gene, which is implicated in hereditary and acquired colorectal cancer and also in some other neoplasms.[9] The function of *APC* is not fully understood. It associates with plakoglobin and β-catenin and may be involved in the regulation of epithelial cell–cell adhesion.[10] Through binding to these proteins, it may also influence a *Wnt-1*-stimulated cell-signalling pathway, controlling the differentiation fate of epithelial cells.[11,12] *In vitro APC* protein also acts as a microtubule binding protein and enhances microtubule polymerization.[13,14] It is unknown if *APC* plays a role in lung cancer and there is no increased incidence of lung cancer in patients with familial adenomatous polyposis (FAP), but we have recently shown that the *APC* protein is normally expressed in respiratory epithelium.[15]

The aims of the present study were to determine the frequency of 5q21 involvement in NSCLC and in particular to determine the relative occurrence of loss in squamous carcinoma and adenocarcinoma, the most common forms of NSCLC, as there is emerging evidence that there may be significant differences in the allelotypes of these neoplasms.[16] These differences may be of practical significance, since epidemiological investigations have shown that the relative incidences of these tumours are changing with adenocarcinoma becoming more prevalent.[17] This shift in the histological profile of lung cancer has led some authors to re-evaluate the WHO classification, which recognizes several morphological subtypes of adenocarcinoma, but has not been shown to be of prognostic or biological significance. It has been suggested that a more clinically relevant distinction can be drawn between adenocarcinomas having origin (like squamous carcinoma and SCLC) from large

Addressee for correspondence: C. A. Cooper, Sir Alastair Currie CRC Laboratories, Molecular Medicine Centre, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, U.K.

Received 24 January 1996
Accepted 29 March 1996

CCC 0022–3417/96/090033–05

airways and those arising in the distal lung parenchyma, from bronchiolar and alveolar lining cells.[18] A comparison between the bronchial and parenchymal groups of adenocarcinomas was therefore included in the study.

Our strategy was to determine the frequency of LOH at four polymorphic loci in *APC* and *MCC* in a series of surgically resected NSCLC cases. LOH at the appropriate genetic locus is not, of itself, sufficient evidence to implicate a particular tumour suppressor gene in the pathogenesis of a neoplasm. The widely accepted multistage model of carcinogenesis predicts that development of neoplasia requires the remaining copy of the gene to be mutated or otherwise inactivated.[9] Mutations in the *APC* gene have been detected in a number of tumours, including those of the colon, pancreas, and stomach.[9,19,20] The majority of mutations have been found in the first half of the gene's coding sequence, two-thirds of somatic mutations in colorectal tumours being restricted to a region called the mutation cluster region (MCR), between codons 1286 and 1513.[21] We therefore decided to use an established single-strand conformational polymorphism (SSCP) technique to screen for mutations in the MCR of the remaining *APC* allele of those cases showing LOH at 5q21.

MATERIALS AND METHODS

Cases and tissue samples

DNA was extracted from paraffin blocks of 97 surgical resections (lobectomy and pneumonectomy) for primary squamous cell carcinoma or adenocarcinoma of the lung, using standard methods.[22] Separate samples were taken from each case to include tumour and non-tumour (usually uninvolved lymph node) tissue. The tumour samples were dissected from 10 μm sections under histological guidance, to minimize stromal/inflammatory cell contamination. Cases included were chosen as follows: 32 consecutive squamous carcinomas, 58 consecutive adenocarcinomas, and 7 additional non-randomly selected adenocarcinomas, all of bronchial type. The latter were included because our experience has shown that bronchial adenocarcinomas are rather less common than those arising in the pulmonary parenchyma.[23] Adenocarcinomas were allocated, on the basis of macroscopic and histopathological appearances, into parenchymal or bronchial subtypes, with an additional category for neoplasms of uncertain or mixed histogenesis. Parenchymal carcinomas were defined by lack of a bronchial origin and by a 'lepidic' growth pattern. Cytologically, they were composed of cells resembling type 2 pneumocytes, clara cells ('tongue-shaped' cells),[17] or mucus cells. The bronchial adenocarcinomas showed obvious origin from and disruption of bronchial mucosa and tended to have an acinar or papillary pattern.

LOH in APC *and* MCC *genes*

LOH for each tumour was determined using polymerase chain reaction (PCR) amplification of intragenic sequences containing known polymorphic sites within the *APC* and *MCC* genes. Restriction fragment length

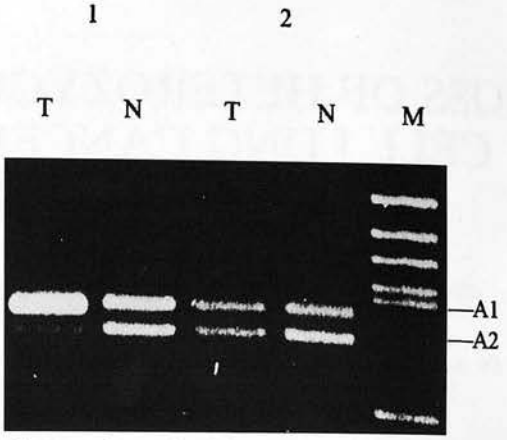


Fig. 1—Two samples analysed by PCR amplification of the naturally-occurring length polymorphism at *MCC* exon 10. In each case, the 'normal' DNA sample (reactive lymph node) is shown in lane N. Both patients are heterozygous for this locus, indicated by A1 (175) and A2 (161) base-pairs. The corresponding tumour samples (T) show loss of A2 in patient 1 (LOH). Heterozygosity is retained in patient 2. Lane M contains molecular weight markers

polymorphisms (RFLPs) in the 3′ untranslated region (UTR)[24] and exon 11 of the *APC* gene[25] and a RFLP in the 3′UTR of *MCC*[26] were analysed using enzymes *Ssp1*, *Rsa1*, and *Mae111*, respectively. A fourth polymorphism, a length polymorphism in *MCC* exon 10,[27] was also analysed. Conditions and primers have been previously published[28] with the exception of a new *APC* 3′ UTR primer. Here the upstream primer was replaced with a new primer (5′ GAAGAGACTGCA ATGTCTAAGAA 3′) to produce a smaller 318 bp PCR product, which was easier to amplify from archival paraffin-embedded tissues. A control DNA fragment containing the appropriate restriction site was added to the *APC* exon 11 (*Rsa1*) and the *MCC* 3′UTR (*Mae111*) digests to ensure complete restriction enzyme cleavage. The digested PCR products were run on agarose or polyacrylamide gels and stained with ethidium bromide. LOH at a given locus was defined by a reduction in the intensity of an allele band of the tumour in comparison with its corresponding normal (Fig. 1). Cases polymorphic at two or more loci were subjected to statistical analysis.

SSCP analysis of the MCR of the APC

All cases showing LOH at 5q21 were screened for *APC* mutations. DNA fragments were amplified by PCR (Table I). Some of the primers used have previously been published.[29] Cycling conditions consisted of an initial denaturation step of 5 min at 94°C, followed by 36 cycles of 30 s denaturation at 94°C, 30 s at the appropriate annealing temperature (see Table I), and 60 s extension at 72°C, with a final extension step of 10 min at 72°C. One microlitre of denaturing solution (0·5 M NaOH, 10 mM EDTA) was added to 5 μl of PCR product. Samples were incubated for 5 min at 50°C, then 3 μl of stop solution [95 per cent formamide, 10 mM EDTA (pH 8), 0·01 per cent bromophenol blue, 0·01 per cent xylene cyanol] was added to each sample following

Table I—Sections of *APC* MCR for SSCP analysis

Nucleotide location	Product size (bp)	Primers 5′→3′	Annealing temperature (°C)
3717–3973 or 3717–3963	256 or 246	AAGTGGTCAGCCTCAAAAGG TGGAACTTCGTCACAGGAT or CTCACAGGATCTTCAGCTGA	58
3870–4164	294	TCAGACGACACAGGAAGCAG GTACATCTGCTAAACATGAGTGGG	56
4132–4423	291	CAGGAGACCCCACTCATGTT CAGCATTTACTGCAGCTTGC	56
4411–4614	203	AGAGTGGACCTAAGCAAGCT CATTTTCCTGAACTGGAGGC	55

incubation. Nine microlitres of sample was loaded onto a non-denaturing MDE gel (AT Biochem, PA, U.S.A.) containing 5 per cent glycerol in 1X TBE and run at 25°C at 25 W. Gels were subsequently silver-stained according to the manufacturer's instructions (Bio-rad Laboratories Ltd., U.K.). Colorectal adenocarcinoma cases previously shown to harbour mutations in the MCR of *APC* were included as positive controls.

RESULTS

APC/MCC *LOH in NSCLC*

A total of 97 NSCLC cases were studied for LOH within the *APC* and *MCC* genes. Sixty-four of the 90 consecutive tumours were informative at two or more loci (41 adenocarcinoma, 23 squamous carcinomas). LOH was seen in 37 per cent (15/41) of adenocarcinomas and 48 per cent (11/23) of squamous carcinomas, giving an overall LOH of 41 per cent in our informative NSCLC cases. Statistical analysis indicated no significant difference in the frequency of LOH between squamous carcinomas and adenocarcinomas ($\chi^2=0{\cdot}772$, $df=1$; $P<0{\cdot}2$).

Within the adenocarcinoma category, which included the additional non-randomly selected bronchial adenocarcinomas, there appeared to be a difference between tumours of bronchial and parenchymal origin, with 5/9 (56 per cent) informative adenocarcinomas arising in the bronchus showing LOH, while parenchymal cancers showed LOH in 6/21 cases (29 per cent). This difference falls short of significance at the 5 per cent level. Sixteen adenocarcinomas could not be clearly described as arising in the bronchus or the parenchyma. LOH was seen in eight (50 per cent) of these tumours.

There was no relationship between LOH and tumour stage. In all cases that were informative for polymorphisms in both genes analysed, loss or retention was always seen in both *APC* and *MCC*; there were no cases showing discordance between the two genes.

Mutation analysis of the APC *gene by SSCP*

Thirty-two lung cancer cases were screened for mutations in the MCR of *APC* by SSCP. This group consisted of 21 adenocarcinomas and 11 squamous carcinomas. The adenocarcinoma group included five non-randomly selected cases of bronchial origin and one case of parenchymal origin which showed LOH at one or more polymorphic loci. In no case was there a difference in the pattern of band mobility between control and tumour DNA, providing strong evidence against the presence of *APC* mutation in any of these NSCLC cases.

DISCUSSION

The long arm of chromosome 5 has recently been the focus of much attention in tumour biology. Particular attention has been paid to the 5q21 region, the locus of *APC*, the gene responsible for FAP.[25] LOH at 5q21 has been described in a variety of neoplasms including carcinomas of the colorectum, stomach, pancreas, oesophagus, and kidney.[19,28,30–32] In the present study, we have shown that LOH at 5q21 is also a feature of some 41 per cent of randomly selected NSCLC cases, a higher percentage than in a recently published study,[8] and that NSCLC exhibiting LOH is not associated with mutation of the recognized MCR of *APC*.

The *APC* gene has previously been shown to exhibit LOH in various tumour types and a significant mutation rate in the MCR of this gene has been described in carcinomas of the colorectum, pancreas, and stomach.[9,19,20] One previous study has looked at a very limited series of lung cancers, finding no evidence of mutation in seven neoplasms showing genetic loss at this locus.[19] Our 32 cases showing LOH were screened using SSCP and no mutations were detected in the MCR. To compute the probability that such a result might have been obtained by chance, we used the following logic. In sporadic colorectal cancer, more than 60 per cent of mutations occur in the MCR, but analysis of these colorectal tumours using the SSCP technique has detected mutations in only 50 per cent of cases.[21] Therefore it appears that SSCP is detecting approximately 80 per cent of mutations. If mutations in the MCR of the *APC* gene were detected as frequently in

NSCLC as in colorectal tumours (50 per cent), then the probability of obtaining a negative result in our SSCP analysis could be calculated by $0{\cdot}5^{32}=2{\cdot}33^{-10}$. Similarly, there is 95 per cent probability that the prevalence of detectable mutations is less than 9 per cent. This is calculated by taking the statistical confidence limit of $P<0{\cdot}05$ and solving for x in the equation $(1-x)^{32}=0{\cdot}05$; $x=0{\cdot}09$). This is strong evidence against the involvement of the gene in pulmonary carcinogenesis, a finding which is consistent with tumour epidemiology, since there is no increase in incidence of lung cancer in FAP patients.

The *MCC* gene at 5q21 has been described as a potentially important tumour suppressor gene. *MCC* was originally identified in colorectal cancer, but more recent mutation analysis by SSCP has failed to show mutations in a series of 80 carcinomas of the large bowel.[28] We know of no reports of *MCC* mutation outside the colorectum and it therefore seems increasingly likely that the status of MCC as a tumour suppressor gene is in some doubt. Further investigation of this association will be necessary, as is a detailed mapping of losses at 5q, to localize any other possible tumour suppressor genes. Interestingly, a recent study has shown that LOH at a novel locus (del 27) lying centromeric to 5q21 is seen in a high proportion of lung cancers.[33]

The few previous reports of genetic losses at 5q21 in lung cancer have largely concentrated on SCLC.[6,7] A recently published study showed LOH in 29 per cent of informative NSCLC cases with a marginal but significant difference in the frequency of LOH between squamous carcinoma and adenocarcinoma, the former being the more frequent.[8] Our larger study suggests a higher incidence of LOH in NSCLC (41 per cent) but no difference in LOH between squamous carcinoma and adenocarcinoma.

Recently there has been accumulating evidence of a shift in the histological pattern of pulmonary neoplasia, with a relative increase in the proportion of adenocarcinomas, particularly when compared with squamous carcinomas and SCLC.[9] The recent rise in the incidence of adenocarcinomas has been noted particularly to involve those arising in the periphery of the lung.[34] Morphological and immunohistochemical studies have shown that there may be biological differences between adenocarcinomas of bronchial compared with those of parenchymal origin. The latter arise on a background of alveolar atypical hyperplasia, with demonstrable increase in expression of the products of the p53 and c-*erb*B-2 genes.[35] It is also evident that adenocarcinomas of bronchial origin have a distinctly poorer clinical prognosis.[18] Our demonstration that LOH at 5q21 appears to be associated more with bronchial than parenchymal adenocarcinomas is intriguing and may, if confirmed in a larger series, provide further biological support for the proposed subclassification of adenocarcinomas. It may be that LOH at 5q21 is associated with the common bronchial neoplasms of all types (squamous carcinoma, adenocarcinoma, and SCLC) and is of lesser importance in the increasingly frequent parenchymal adenocarcinoma.

We conclude that LOH at 5q21 is more common in SCLC than in NSCLC, with no difference in the frequency of LOH between adenocarcinomas and squamous carcinomas. LOH was detected more frequently in adenocarcinomas of parenchymal rather than bronchial origin, although these results do not reach statistical significance. Finally, it seems unlikely that *APC* plays an important role in NSCLC.

ACKNOWLEDGEMENT

This study was supported by a grant from the Scottish Office Home and Health Department.

REFERENCES

1. Dunnill MS, Gatter KC. Cellular heterogeneity in lung cancer. *Histopathology* 1986; **10:** 461–475.
2. World Health Organisation. Histological Typing of Lung Tumours. 2nd edn. Geneva: WHO, 1981.
3. Wong AJ, Ruppert JM, Eggleton J, Hamilton SR, Baylin SB, Vogelstein B. Gene amplification of c-*myc* and n-*myc* in small cell carcinoma of the lung. *Science* 1986; **23:** 461–464.
4. Tateishi M, Ishida T, Mitsudomi T, Kaneko S, Sugimachi K. Prognostic value of c-*erb*B-2 protein expression in human lung adenocarcinoma and squamous cell carcinoma. *Eur J Cancer* 1991; **27:** 1372–1375.
5. Ashton-Rickardt PG, Wyllie AH, Bird CC, *et al. MCC*, a candidate familial polyposis gene in 5q21, shows frequent allele loss in colorectal and lung cancer. *Oncogene* 1991; **6:** 1881–1886.
6. Miura I, Graziano SL, Cheng JQ, Doyle LA, Testa JR. Chromosome alterations in human small cell lung cancer: frequent involvement of 5q. *Cancer Res* 1991; **52:** 1322–1328.
7. D'Amico D, Carbone DP, Johnson BE, Meltzer SJ, Minna JD. Polymorphic sites within the *MCC* and *APC* loci reveal very frequent loss of heterozygozity in human small cell lung cancer. *Cancer Res* 1992; **52:** 1996–1999.
8. Fong KM, Zimmerman PV, Smith PJ. Tumor progression and loss of heterozygosity at 5q and 18q in non-small cell lung cancer. *Cancer Res* 1995; **55:** 220–223.
9. Bishop DT, Hall NR. The genetics of colorectal cancer. *Eur J Cancer* 1994; **30:** 1946–1956.
10. Munemitsu S, Albert I, Souza B, Rubinfeld B, Polakis P. Regulation of intracellular β-catenin levels by the adenomatous polyposis coli (*APC*) tumor-suppressor protein. *Proc Natl Acad Sci USA* 1995; **92:** 3046–3050.
11. Hinck L, Nelson W, Papkoff J. *Wnt-1* modulates cell–cell adhesion in mammalian cells stabilising β-catenin binding to the cell adhesion protein cadherin. *J Cell Biol* 1994; **124:** 729–741.
12. Funayama N, Fagotto F, McCrea P, Gumbiner B. Embryonic axis induction by the armidillo repeat domain of β-catenin: evidence for intracellular signalling. *J Cell Biol* 1995; **128:** 959–968.
13. Munemitsu S, Souza B, Müller O, Albert I, Rubinfeld B, Polakis P. The APC gene product associates with microtubules *in vivo* and promotes their assembly *in vitro*. *Cancer Res* 1994; **54:** 3676–3681.
14. Smith KJ, Levy DB, Maupin P, Pollard TD, Vogelstein B, Kinzler KW. Wild-type but not mutant APC associates with the microtubule cytoskeleton. *Cancer Res* 1994; **54:** 3672–3675.
15. Midgley CA, White S, Howitt R, *et al. APC* expression in normal human tissues. (Submitted).
16. Tsuchiya E, Nakamura Y, Weng SY, *et al.* Allelotype of non-small lung carcinoma—comparison between loss of heterozygosity in squamous cell carcinoma and adenocarcinoma. *Cancer Res* 1992; **52:** 2478–2481.
17. Gilliland FD, Samet JM. Lung cancer. In: Doll R, Fraumeni JF, Muir CF, eds. Trends in Cancer Incidence and Mortality. New York: Cold Spring Harbor Laboratory Press, 1994; 175–195.
18. Edwards CW. Pulmonary adenocarcinoma: a review of 106 cases and proposed new classification. *J Clin Pathol* 1987; **40:** 125–135.
19. Horii A, Nakatsuru S, Miyoshi Y, *et al.* Frequent somatic mutations of the *APC* gene in human pancreatic cancer. *Cancer Res* 1992; **52:** 6696–6698.
20. Nakatsuru S, Yanagisawa A, Ichii S, *et al.* Somatic mutations of the *APC* gene in gastric cancer: frequent mutations in very well differentiated adenocarcinoma and signet-ring cell carcinoma. *Hum Mol Genet* 1992; **1:** 559–563.
21. Miyoshi Y, Nagase H, Ando H, *et al.* Somatic mutations of the *APC* gene in colorectal tumours: mutation cluster region in the *APC* gene. *Hum Mol Genet* 1992; **4:** 229–233.
22. Levi S, Urbono-Ispizua A, Gill R, *et al.* Multiple K-*ras* codon 12 mutations in cholangiocarcinomas demonstrated with a sensitive polymerase chain reaction technique. *Cancer Res* 1991; **51:** 3497–3502.
23. Carey FA, Wallace WAH, Fergusson RJ, Kerr KM, Lamb D. Alveolar atypical hyperplasia in association with primary pulmonary adenocarcinoma: a clinico-pathological study of 10 cases. *Thorax* 1992; **47:** 1041–1043.

24. Heighway J, Hoban PR, Wyllie AH. *Ssp1* polymorphism in sequence encoding 3′ untranslated region of the *APC* gene. *Nucleic Acids Res* 1991; **19:** 6966.
25. Groden J, Thilveris A, Samowitz W, *et al.* Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* 1991; **66:** 589–600.
26. Heighway J. Hoban PR, Wyllie AH. *Mae111* polymorphism in sequence encoding 3′ untranslated region of the *MCC* gene. *Nucleic Acids Res* 1992; **20:** 1160.
27. Miyoshi Y, Nishisho I, Miki Y, *et al.* Insertion/deletion polymorphism and other restriction fragment length polymorphisms in the *MCC* gene. *Jpn J Cancer Res* 1992; **83:** 10–14.
28. Curtis LJ, Bubb VJ, Gledhill S, Morris RG, Bird CC, Wyllie AH. Loss of heterozygosity of *MCC* is not associated with mutation of the retained allele in sporadic colorectal cancer. *Hum Mol Genet* 1994; **3:** 443–446.
29. Curtis L, Wyllie AH, Shaw J, *et al.* Evidence against involvement of *APC* mutations in papillary thyroid carcinoma. *Eur J Cancer* 1994; **7:** 984–987.
30. Sano T, Tsujino T, Yoshida K, *et al.* Frequent loss of heterozygosity on chromosome 1q, 5q and 17p in human gastric carcinomas. *Cancer Res* 1991; **51:** 2926–2931.
31. Boynton RF, Blount PL, Yin J, *et al.* Loss of heterozygosity involving the *APC* and *MCC* genetic loci occurs in the majority of human esophageal cancers. *Proc Natl Acad Sci USA* 1992; **89:** 3385–3388.
32. Morita R, Saito S, Ishikawa J, Ogawa O, Yoshida O, Yamakawa K. Common regions of deletion on chromosomes 5q, 6q and 10q in renal cell carcinoma. *Cancer Res* 1991; **51:** 5817–5820.
33. Wieland I, Bohm M. Frequent allelic deletion at a novel locus on chromosome 5 in human lung cancer. *Cancer Res* 1994; **54:** 1772–1774.
34. Auerbach O, Garfinkel L. The changing pattern of lung carcinoma. *Cancer* 1991; **68:** 1973–1977.
35. Kerr KM, Carey FA, King G, Lamb D. Atypical alveolar hyperplasia: relationship with pulmonary adenocarcinoma, p53 and c-*erb*B-2 expression. *J Pathol* 1994; **174:** 249–256.

JOURNAL OF PATHOLOGY, VOL. 181: 401–404 (1997)

THE PATTERN OF K-*ras* MUTATION IN PULMONARY ADENOCARCINOMA DEFINES A NEW PATHWAY OF TUMOUR DEVELOPMENT IN THE HUMAN LUNG

CINDY A. COOPER[1], FRANCIS A. CAREY[2*], VIVIEN J. BUBB[1], DAVID LAMB[1], KEITH M. KERR[3] AND ANDREW H. WYLLIE[1]

[1]*Sir Alastair Currie CRC Laboratories, Department of Pathology, University of Edinburgh, Edinburgh, U.K.,* [2]*Department of Pathology, Ninewells Hospital and Medical School, Dundee, U.K.,* [3]*Department of Pathology, Aberdeen Royal Infirmary, Aberdeen, U.K.*

SUMMARY

Codon 12 of the K-*ras* oncogene was screened for mutations in 65 surgically-resected primary pulmonary adenocarcinomas and in 32 tissue foci of alveolar atypical hyperplasia (AAH) by a polymerase chain reaction (PCR)-based method. Mutations in either position 1 or position 2 of codon 12 were detected in 16 tumours (25 per cent). When analysed by site of origin, mutations were seen in 9/26 (35 per cent) parenchymal and in 0/12 bronchial adenocarcinomas ($P<0{\cdot}02$). K-*ras* mutations were seen in five AAH lesions from four patients. DNA sequencing showed that the great majority of mutations in both adenocarcinomas and AAH were G–T transversions. These findings provide support for the classification of pulmonary adenocarcinomas into bronchial and parenchymal subtypes and also provide molecular evidence to support the importance of AAH in the development of parenchymal cancers. © **1997 by John Wiley & Sons, Ltd.**

J. Pathol. 181: 401–404, 1997.
No. of Figures: 2. No. of Tables: 1. No. of References: 17.

KEY WORDS—**lung neoplasms; adenocarcinoma; K-*ras*; alveolar atypical hyperplasia**

INTRODUCTION

Important epidemiological and biological features distinguish adenocarcinoma of the lung from other common malignancies arising at this site.[1] Little is known for certain concerning the histogenesis of adenocarcinoma, however, and the morphological classification is somewhat controversial. The World Health Organisation classification of lung cancer recognizes several histological subtypes of primary pulmonary adenocarcinoma but has been widely criticized as poorly reproducible and lacking a clinical or biological basis.[2,3] One alternative classification is based on recognition of the differences between those adenocarcinomas taking origin (like squamous and small cell tumours) from the bronchi and those (the majority) developing in the distal parenchyma.[3] We have previously proposed that parenchymal adenocarcinomas may arise on a background of alveolar atypical hyperplasia (AAH).[4] These lesions are characterized by distinct cytological atypia and have also recently been described in other populations, particularly in North America and Japan.[5,6] AAH has been shown to have abnormal cytometric DNA profiles and to show immunohistochemical evidence of abnormal expression of the p53 and c-*erb*-B2 gene products.[7,8]

Mutations in codon 12 of the K-*ras* oncogene have been shown to occur in between approximately 10 and 40 per cent of pulmonary adenocarcinomas.[9–11] It has been suggested that the presence of mutation may be a marker of an adverse prognosis in these neoplasms.[10,11] One study in a Japanese population has suggested that K-*ras* mutations are associated with tumour differentiation, being seen more often in mucinous than in other types of adenocarcinoma.[12] It is, as yet, unclear whether there are differences in the occurrence of mutation between bronchial and parenchymal adenocarcinomas.

The aims of the present study were to determine whether the K-*ras* mutation associates predominantly with one or other subtype and to investigate the occurrence of these mutations in the potentially premalignant lesions of AAH.

MATERIALS AND METHODS

Tissue samples

All cases included in this study have previously been reported in an investigation of loss of heterozygosity at chromosome 5q21 in lung cancer.[13] DNA was extracted from tumour-containing paraffin blocks of 65 surgically resected primary adenocarcinomas of lung using previously described methods.[13] Tumour samples were selected so as to minimize stromal and inflammatory contamination. Non-tumour (control) DNA was extracted from uninvolved lymph node tissue. The series of 65 comprised 58 consecutive resected adenocarcinomas, 26 (45 per cent) being of parenchymal type, 5 (9 per cent) showing a bronchial origin, and 27 (47 per cent) that could not be accurately classified. Seven additional bronchial adenocarcinomas were included as these lesions are, in our experience, somewhat less common

*Correspondence to: Dr F. A. Carey, Department of Pathology, Ninewells Hospital and Medical School, Dundee DD1 9SY, U.K.

Contact grant sponsor: Chief Scientist's Office, Scottish Home and Health Department.

CCC 0022–3417/97/040401–04 $17.50

Received 24 July 1996
Accepted 4 October 1996

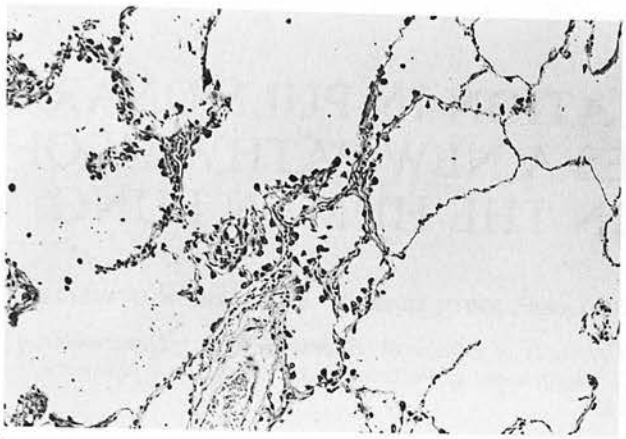


Fig. 1—A focus of AAH. The normal pulmonary parenchymal architecture is retained but shows alveolar wall thickening. The lesion is lined by a population of plump, cytologically atypical cells morphologically resembling type II alveolar lining cells. This particular lesion measured 2·5 mm in maximum dimension and harboured a G–T transversion at position 1 of codon 12 of K-*ras*

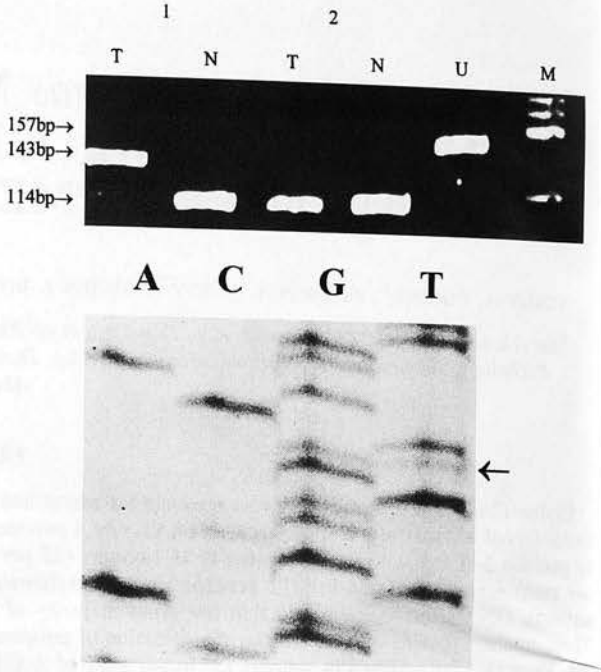


Fig. 2—Detection and sequencing of K-*ras* mutations. (a) Paired tumour (T) and normal (N) samples are run against uncut PCR product (U) and molecular weight markers (M). Case 1 shows complete cleavage of the tumour DNA but with loss of the *Bst*N1 restriction site extending into codon 12, thus generating a 'mutated' band of 143 bp. Case 2 shows no evidence of mutation. (b) Sequencing gel showing a G–T transversion at position 1 of codon 12 (arrowed)

than parenchymal carcinomas.[13] Classification of the adenocarcinomas into bronchial and parenchymal subtypes was based on review of macroscopic and histological features. The parenchymal tumours were defined by a lipidic growth pattern and contained significant populations of cells resembling type 2 alveolar lining cells or Clara cells.[3] Bronchial adenocarcinomas were centred on, and frequently disrupted, large airways and tended to show an acinar or papillary pattern. Tumours not clearly fitting either of these categories, often larger, poorly differentiated neoplasms, were considered separately as of uncertain or mixed histogenesis.

AAH was sought in non-tumour parenchyma of all cases included in this study (Fig. 1). Alveolar cell hyperplasia or atypia found in direct continuity with obvious tumour was not classified as AAH. Material used in a previous investigation was reviewed and included where sufficient AAH tissue was available.[8] Sections of the AAH were cut at 10 μm, dewaxed, and stained with eosin. The slides were examined under the microscope and areas of AAH marked and removed by microdissection. DNA was extracted into a volume of 50–100 μl of lysis buffer, depending on the size of the AAH lesion. The lysis buffer comprised 10 mM Tris at pH 8·3, 50 mM KCl, 0·45 per cent (v/v) Tween 20, and 2·5 mM $MgCl_2$ containing 1·25 mg/ml proteinase K.

*Screening for K-*ras *codon 12 mutations*

Polymerase chain reaction (PCR) was applied to tumour and normal DNA to amplify a 157 base pair (bp) product encompassing codon 12 of the K-*ras* gene. The primers used incorporated mismatches designed to create two *Bst*N1 restriction endonuclease sites in the normal gene.[14] One of these cleavage sites was within the 3′ primer; it acted as a control for complete *Bst*N1 digestion and generated a 14 bp fragment. The 5′ primer, within exon 1, contained a cytosine substitution corresponding to the first position of codon 11, thus creating a *Bst*N1 recognition site extending into positions 1 and 2 of codon 12 which are outside the primer sequences. Mutations in either of the first two positions of codon 12 result in loss of this restriction site and can thus be identified by gel electrophoresis.

PCRs were carried out in 100 μl volumes containing 10 μl of DNA, 0·5 μM of each primer, 1·5 mM $MgCl_2$, 0·2 mM of each dNTP, and 2·5 units of Taq polymerase (Gibco BRL, U.K.) made up in PCR buffer (Gibco BRL, U.K.). Cycling conditions were as follows: an initial denaturation step of 5 min at 94°C followed by 30–38 cycles of 30 s denaturation at 94°C, annealing at 55°C, and extension at 72°C, with a final extension step of 10 min at 72°C. Following amplification, 20 μl of PCR product was digested at 60°C overnight with 10 units of *Bst*NI [New England Biolabs (U.K.) Ltd.]. Digested PCR products were run on a 3 per cent agarose gel stained with ethidium bromide and sizes compared with molecular weight markers. Samples which did not contain codon 12 mutations generated 114 bp fragments whilst mutated samples, having lost the second *Bst*N1 restriction site, yielded larger fragments at 143 bp (Fig. 2).

*Sequencing K-*ras *mutations*

Samples shown to carry a mutation at codon 12 were analysed by sequencing. Tumour and normal samples were amplified by PCR using a biotinylated exon 1 primer 5′GACTGAATATAAACTTGTGG3′ and the previous intron 1 3′ primer. Cycling conditions were as stated above. Single-stranded DNA was isolated using

Table I—K-*ras* mutational analysis in alveolar atypical hyperplasia (AAH) in 16 patients. K-*ras* mut refers to the presence or absence of abnormal bands on *Bst*N1 digests. The normal codon 12 sequence is GGT

Case No.	AAH	K-*ras* mut	Sequence
1	1a	+	GAT
	1b	+	TTT
	1c	−	~
	1d	−	~
	1e	−	~
2	2a	−	~
	2b	−	~
3	3a	−	~
	3b	−	~
	3c	−	~
4	4a	−	~
	4b	−	~
5	5a	−	~
6	6a	−	~
	6b	−	~
	6c	−	~
7	7a	−	~
	7b	+	GTT
	7c	−	~
	7d	−	~
	7e	−	~
	7f	−	~
8	8a	−	~
9	9a	−	~
10	10a	−	~
	10b	−	~
11	11a	−	~
12	12a	−	~
13	13a	+	TGT
14	14a	−	~
15	15a	+	TGT
16	16a	−	~

streptavidin-coated magnetic beads (Dynal, Norway) and sequenced using the Sequenase version II sequencing kit (Amersham International, U.K.).

In some AAH lesions, possibly because of a high 'background' of wild-type K-*ras*, mutations were only indistinctly visible on both the *Bst*N1 and the sequencing gels. In these cases (1a, 1b, 15a of Table I) a further PCR was included to enrich for the mutant allele. *Bst*N1-digested samples were amplified using the biotinylated exon 1 primer and a new downstream primer in exon 1 (5′CTCTATTGTTGGATCATATT 3′). Amplified samples were then sequenced as before. The unusual pattern of mutation seen in lesions 1a and 1b (Table I) was confirmed by cloning the PCR product (Original TA Cloning Kit, Invitrogen) and sequencing again as before.

RESULTS

The *Bst*N1 assay for K-*ras* mutation was successfully applied to all 65 adenocarcinomas. Mutations in codon 12 of the gene were detected in 16 cases (25 per cent). DNA sequencing showed that nine tumours had G–T transversions at position 1; G–T transversions at position 2 were seen in six cases; and one had a G–A transition at position 2 (Fig. 2). There was a strong relationship between tumour subtype and presence of K-*ras* mutation, with 9/26 (35 per cent) parenchymal lesions showing mutation while none of the 12 bronchial adenocarcinomas had evidence of codon 12 abnormality ($P=0{\cdot}019$, Fisher's exact test). The mutation incidence in the unclassified tumours (7/27, 26 per cent) was the same as for the series as a whole.

Foci of AAH were available from 16 patients yielding a total of 32 lesions, with between one and six distinct foci per case (mean 2·0). The results of the mutational analysis on these lesions are summarized in Table I. Evidence of K-*ras* mutation was seen in five AAH (15 per cent) from four patients. One of these patients had an associated adenocarcinoma. This was not mutated. AAH mutations were confirmed by sequencing (Table I) showing in general a similar pattern to adenocarcinomas but with one unusual lesion showing G–T transversions at both positions 1 and 2 of codon 12. This mutation was confirmed by cloning.

DISCUSSION

Classification of primary pulmonary adenocarcinomas into parenchymal and bronchial subtypes has been advocated as a clinically relevant distinction, with the latter having a worse prognosis.[3] Histologically, bronchial adenocarcinomas share morphological features with either the normal bronchial lining cells or the cells of bronchial glands, while parenchymal cancers frequently show alveolar lining or Clara cell differentiation.[3] In a previous study, we have shown that loss of heterozygosity at chromosome 5q21, although occurring in a proportion of both tumour types, appears to be more common in tumours of bronchial than parenchymal origin.[13] Here we have analysed the same tumours and have demonstrated a significant difference in the frequency of K-*ras* codon 12 mutations between the two subtypes, K-*ras* mutations being present in about one-third of adenocarcinomas arising in the parenchyma but not detected in bronchial adenocarcinomas. We have compared the results of the studies and have found no evidence of a relationship between K-*ras* mutation status and retention or loss of heterozygosity at 5q. There is now clear evidence for a genetic difference between the two subtypes of adenocarcinoma. Although we believe this to be the first unequivocal demonstration of this difference, it is interesting to note that, in other reported studies in lung tumours, the incidence of K-*ras* codon 12 mutations varies from 10 to 40 per cent, perhaps reflecting different ratios of bronchial to parenchymal adenocarcinomas included in these studies.[9–12]

The data also support the proposal that parenchymal (but not bronchial) adenocarcinomas derive from AAH, since this shares with parenchymal adenocarcinoma a predilection for clonal expansion of cells bearing

mutated K-*ras*. Foci of hyperplasia of atypical alveolar lining cells, usually with an associated expansion of the interstitial connective tissue, have been noted for at least 30 years but it is only recently that the significance of these lesions as putative progenitors of parenchymal adenocarcinoma has been emphasized.[4,8,15] Initial suspicion that AAH is a premalignant phenotype was on morphological criteria but this has been heightened by immunohistochemical evidence of expression of proliferation antigens and of the c-*erb*-B2 and stabilized p53 proteins.[4,8] Our demonstration of mutations in codon 12 of the K-*ras* oncogene is unequivocal molecular proof of genetic abnormality in AAH and confirms a recently published study describing K-*ras* mutations in 16/41 AAH.[16]

Finally, the nature of the mutations in K-*ras* gives some clue to the underlying carcinogenic process in AAH and parenchymal adenocarcinomas of the lung. Sequence analysis revealed that 94 per cent of mutations were G–T transversions, a predominance reflected in other point mutations in cancer-related genes in respiratory epithelium.[17] This type of mutation is entirely consistent with origin from mismatched nucleotides inserted in attempted repair of deoxyguanosine oxidation metabolites, presumably generated by exposure to inhaled carcinogens. Our data also clearly show that separate foci of AAH within the lung are genetically distinct with respect to K-*ras* mutation. Several patients had both mutated and non-mutated AAH (Table I). One had a mutated AAH but no evidence of mutation in the co-existing adenocarcinoma. Interestingly, one individual (patient 1, Table I) harboured two AAH lesions with different K-*ras* mutations, one of which was a double G–T transversion involving positions 1 and 2 of codon 12. These associations would tend to refute early speculation that AAH lesions might possibly represent intrapulmonary spread of primary adenocarcinoma and further strengthen the putative role of AAH as an independent neoplastic proliferation and a possible precursor of parenchymal adenocarcinoma.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Chief Scientist's Office, Scottish Home and Health Department.

REFERENCES

1. Auerbach O, Garfinkel L. The changing pattern of lung carcinoma. *Cancer* 1991; **68:** 1973–1977.
2. World Health Organisation. Histological Typing of Lung Tumours. 2nd edn. Geneva: WHO, 1981.
3. Edwards CW. Pulmonary adenocarcinoma: review of 106 cases and proposed new classification. *J Clin Pathol* 1987; **40:** 125–135.
4. Carey FA, Wallace WAH, Fergusson RJ, Kerr KM, Lamb D. Alveolar atypical hyperplasia in association with primary pulmonary adenocarcinoma: a clinicopathological study of 10 cases. *Thorax* 1992; **47:** 1041–1043.
5. Miller RR. Bronchioloalveolar cell adenomas. *Am J Surg Pathol* 1990; **14:** 904–912.
6. Weng S, Tsuchiya E, Satoh Y, *et al.* Multiple atypical adenomatous hyperplasia of type II pneumocytes and bronchiolo-alveolar carcinoma. *Histopathology* 1990; **16:** 101–103.
7. Nakayama H, Noguschi M, Tsuchiya R, Kodoma T, Shimosato Y. Clonal growth of atypical adenomatous hyperplasia of the lung: cytofluorometric analysis of nuclear DNA content. *Mod Pathol* 1990; **3:** 314–320.
8. Kerr KM, Carey FA, King G, Lamb D. Atypical alveolar hyperplasia: relationship with pulmonary adenocarcinoma, p53 and c-erb B-2 expression. *J Pathol* 1994; **174:** 249–256.
9. Rodenhuis S, Van de Wetering ML, Mooi WJ, Evers SG, van Zandwijk N, Bos JL. Mutational activation the K-*ras* oncogene: a possible pathogenetic factor in adenocarcinoma of the lung. *N Engl J Med* 1987; **317:** 929–935.
10. Slebos RJC, Kibbelaar RE, Dalesio O, *et al.* K-*ras* oncogene activation as a prognostic marker in adenocarcinoma of the lung. *N Engl J Med* 1990; **323:** 561–565.
11. Mitsudomi T, Steinberg SM, Oie HK, *et al. ras* gene mutations in non-small cell lung cancers are associated with shortened survival irrespective of treatment intent. *Cancer Res* 1991; **51:** 4999–5002.
12. Kobayashi T, Tsuda H, Noguchi M, *et al.* Association of point mutation in c-Ki-*ras* oncogene in lung adenocarcinoma with particular reference to cytologic subtypes. *Cancer* 1990; **66:** 289–294.
13. Cooper CA, Bubb VJ, Smithson N, *et al.* Loss of heterozygosity at 5q21 in non-small cell lung cancer: a frequent event but without evidence of APC mutation. *J Pathol* 1996; **180:** 33–37.
14. Jiang W, Kahn SM, Guillem JG, Lu SH, Weinstein IB. Rapid detection of *ras* oncogenes in human tumours: applications to colon, esophageal and gastric cancer. *Oncogene* 1989; **4:** 923–928.
15. Meyer EC, Liebow AA. Relationship of interstitial pneumonia, honeycombing and atypical epithelial proliferation to cancer of the lung. *Cancer* 1965; **18:** 322–351.
16. Westra WH, Baas IO, Hruban RH, *et al.* K-*ras* oncogene activation in atypical alveolar hyperplasias of the human lung. *Cancer Res* 1996; **56:** 2224–2226.
17. Takeshima Y, Seyama T, Bennett WP, *et al.* p53 mutations in lung cancers from non-smoking atomic bomb survivors. *Lancet* 1993; **342:** 1520–1521.